Quenched Normal Approximation for Random Sequences of Transformations

Hella, Olli

2020-01

Hella , O & Stenlund , M 2020 , ' Quenched Normal Approximation for Random Sequences of Transformations ' , Journal of Statistical Physics , vol. 178 , no. 1 , pp. 1-37 . https://doi.org/10.1007/s10955-019-0

http://hdl.handle.net/10138/311631
https://doi.org/10.1007/s10955-019-02390-5

unspecified
publishedVersion

Journal of Statistical Physics (2020) 178:1–37
https://doi.org/10.1007/s10955-019-02390-5

Quenched Normal Approximation for Random Sequences of Transformations

Olli Hella[1] · Mikko Stenlund[1]

Received: 29 April 2019 / Accepted: 18 September 2019 / Published online: 20 November 2019

Abstract
We study random compositions of transformations having certain uniform fiberwise properties and prove bounds which in combination with other results yield a quenched central limit theorem equipped with a convergence rate, also in the multivariate case, assuming fiberwise centering. For the most part we work with non-stationary randomness and non-invariant, non-product measures. Independently, we believe our work sheds light on the mechanisms that make quenched central limit theorems work, by dissecting the problem into three separate parts.

Keywords Quenched normal approximation · Dynamical systems

Mathematics Subject Classification 60F05 · 37A05 · 37A50

1 The Problem

In the following we will study random compositions $T_{\omega_n} \circ \cdots \circ T_{\omega_1}$ of maps where $\omega = (\omega_n)_{n\ge 1}$ is a sequence drawn randomly from a probability space $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_0^{\mathbb{Z}_+}, \mathcal{E}^{\mathbb{Z}_+}, \mathbb{P})$. Here $(\Omega_0, \mathcal{E})$ is a measurable space and $\mathbb{Z}_+ = \{1, 2, \ldots\}$. For each $\omega_0 \in \Omega_0$, $T_{\omega_0} : X \to X$ is a measurable self-map on the same measurable space $(X, \mathcal{B})$. Consider the shift transformation

$$\tau : \Omega \to \Omega : \omega = (\omega_1, \omega_2, \ldots) \mapsto \tau\omega = (\omega_2, \omega_3, \ldots).$$

We assume that τ is $\mathcal{F}$-measurable, but does not necessarily preserve the probability measure $\mathbb{P}$. Next, define the map

Communicated by Alessandro Giuliani.

✉ Mikko Stenlund
mikko.stenlund@helsinki.fi
http://www.helsinki.fi/~stenlund/

Olli Hella
olli.hella@helsinki.fi

1 Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014 Helsinki, Finland

$$\varphi : \mathbb{N} \times \Omega \times X \to \mathbb{N} \times \Omega \times X : \varphi(n, \omega, x) = T_{\omega_n} \circ \cdots \circ T_{\omega_1}(x) \tag{1}$$

with the convention $\varphi(0, \omega, x) = x$. We assume that the map $\varphi(n, \cdot, \cdot)$ is measurable from $\mathcal{F} \otimes \mathcal{B}$ to $\mathcal{B}$ for every $n \in \mathbb{N} = \{0, 1, \dots\}$. The maps $\varphi(n, \omega) = \varphi(n, \omega, \cdot) : X \to X$ form a cocycle over the shift τ, which means that the identities $\varphi(0, \omega) = \mathrm{id}_X$ and $\varphi(n + m, \omega) = \varphi(n, \tau^m\omega) \circ \varphi(m, \omega)$ hold.

Remark 1.1 There is no fundamental reason for working with one-sided time other than that the randomness in our paper is mostly non-stationary—a context in which the concept of an infinite past is perhaps unnatural. For stationary randomness there is no obstacle for two-sided time. The other reason is plain philosophy: our concern will be the future, and whether the observed system has been running before time 0 we choose to ignore—without damage as long as our assumptions (specified later) hold from time 0 onward.

Consider an observable $f : X \to \mathbb{R}$. Introducing notations, we write

$$f_i = f \circ T_{\omega_i} \circ \cdots \circ T_{\omega_1} = f \circ \varphi(i, \omega)$$

as well as

$$S_n = \sum_{i=0}^{n-1} f_i \quad \text{and} \quad W_n = \frac{S_n}{\sqrt{n}}.$$

Given an initial probability measure μ, we write $\bar{f}_i$ and $\bar{W}_n$ for the corresponding fiberwise-centered random variables:

$$\bar{f}_i = f_i - \mu(f_i) \quad \text{and} \quad \bar{W}_n = W_n - \mu(W_n).$$

Note that all of these depend on ω. Next, we define

$$\sigma_n^2 = \mathrm{Var}_\mu \bar{W}_n = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mu(\bar{f}_i \bar{f}_j).$$

Note that σ_n^2 depends on ω.

It is said that a *quenched* CLT equipped with a rate of convergence holds if there exists $\sigma > 0$ such that $d(\bar{W}_n, \sigma Z)$ tends to zero with some (in our case, uniform) rate *for almost every* ω. Here $Z \sim \mathcal{N}(0, 1)$ and the limit variance σ^2 is *independent* of ω. Moreover, d is a distance of probability distributions which we assume to satisfy

$$d(\bar{W}_n, \sigma Z) \le d(\bar{W}_n, \sigma_n Z) + d(\sigma_n Z, \sigma Z)$$

and

$$d(\sigma_n Z, \sigma Z) \le C|\sigma_n - \sigma|,$$

at least when $\sigma > 0$ and σ_n is close to σ; and that $d(\bar{W}_n, \sigma Z) \to 0$ implies weak convergence of $\bar{W}_n$ to $\mathcal{N}(0, \sigma^2)$. One can find results in the recent literature that allow to bound $d(\bar{W}_n, \sigma_n Z)$; see Nicol–Török–Vaienti [19] and Hella [13]. In this paper we supplement those by providing conditions which allow to identify a non-random σ and to obtain a bound on $|\sigma_n(\omega) - \sigma|$ which tends to zero at a certain rate for almost every ω, which is a key feature of quenched CLTs.

Our strategy is to find conditions such that $\sigma_n^2(\omega)$ converges almost surely to

$$\sigma^2 = \lim_{n\to\infty} \mathbb{E}\sigma_n^2.$$

This is motivated by two observations: (i) if $\lim_{n\to\infty} \sigma_n^2 = \sigma^2$ almost surely, dominated convergence should yield the equation above, and (ii) $\mathbb{E}\sigma_n^2$ is the variance of $\bar{W}_n$ with respect to the product measure $\mathbb{P} \otimes \mu$, since $\mu(\bar{W}_n) = 0$:

$$\mathbb{E}\sigma_n^2 = \mathbb{E}\,\mathrm{Var}_\mu \bar{W}_n = \mathbb{E}\mu(\bar{W}_n^2) = \mathrm{Var}_{\mathbb{P}\otimes\mu} \bar{W}_n.$$

Remark 1.2 One has to be careful and note that $\bar{W}_n$ has been centered fiberwise, with respect to μ instead of the product measure. Therefore, $\mathrm{Var}_{\mathbb{P}\otimes\mu} \bar{W}_n$ and $\mathrm{Var}_{\mathbb{P}\otimes\mu} W_n$ differ by $\mathrm{Var}_{\mathbb{P}}\, \mu(W_n)$:

$$\mathbb{E}\sigma_n^2 = \mathrm{Var}_{\mathbb{P}\otimes\mu} \bar{W}_n = \mathbb{E}\mu(\bar{W}_n^2) = \mathbb{E}\,\mathrm{Var}_\mu \bar{W}_n = \mathbb{E}\,\mathrm{Var}_\mu W_n = \mathrm{Var}_{\mathbb{P}\otimes\mu} W_n - \mathrm{Var}_{\mathbb{P}}\, \mu(W_n).$$

In special cases it may happen that $\mathrm{Var}_{\mathbb{P}}\, \mu(W_n) \to 0$, or even $\mathrm{Var}_{\mathbb{P}}\, \mu(W_n) = 0$ if all the maps T_{ω_i} preserve the measure μ, whereby the distinction vanishes and the use of a non-random centering becomes feasible. We will briefly return to this point in Remark C.2 motivated by a result in [1]. A related observation is made in Remark A.3 which answers a question raised in [2] concerning the trick of "doubling the dimension".

To implement the strategy, we handle the terms on the right side of

$$|\sigma_n^2(\omega) - \sigma^2| \le |\sigma_n^2(\omega) - \mathbb{E}\sigma_n^2| + |\mathbb{E}\sigma_n^2 - \sigma^2|$$

separately, obtaining convergence rates for both. Note that these are of fundamentally different type: the first one concerns almost sure deviations of σ_n^2 about the mean, while the second one concerns convergence of said mean together with the identification of the limit.

Remark 1.3 That the required bounds can be obtained illuminates the following pathway to a quenched central limit theorem:

(1) $d(\bar{W}_n, \sigma_n Z) \to 0$ almost surely,
(2) $\sigma_n^2 - \mathbb{E}\sigma_n^2 \to 0$ almost surely,
(3) $\mathbb{E}\sigma_n^2 \to \sigma^2$ for some $\sigma^2 > 0$,

where the last step involves the identification of σ^2.

Remark 1.4 Let us emphasize that in general we do not assume $\mathbb{P}$ to be stationary or of product form; μ to be invariant for any of the maps T_{ω_i}; or $\mathbb{P} \otimes \mu$ (or any other measure of similar product form) to be invariant for the random dynamical system (RDS) associated to the cocycle φ.

Quenched limit theorems for RDSs are abundant in the literature, going back at least to Kifer [14]. Nevertheless they remain a lively topic of research to date: Recent central limit theorems and invariance principles in such a setting include Ayyer–Liverani–Stenlund [4], Nandori–Szasz–Varju [18], Aimino–Nicol–Vaienti [2], Abdelkader–Aimino [1], Nicol–Török–Vaienti [19], Dragičević et al. [9,10], and Chen–Yang–Zhang [8]. Moreover, Bahsoun et al. [5–7] establish important optimal quenched correlation bounds with applications to limit results, and Freitas–Freitas–Vaienti [11] establish interesting extreme value laws which have attracted plenty of attention during the past years.

Structure of the paper the main result of our paper is Theorem 4.1 in Sect. 4. It is an immediate corollary of Theorem 2.14 of Sect. 2, which concerns $|\sigma_n^2(\omega) - \mathbb{E}\sigma_n^2|$, and of Theorem 3.9 of Sect. 3, which concerns $|\mathbb{E}\sigma_n^2 - \sigma^2|$. In Sect. 4 we also explain how the results of this paper extend to the vector-valued case $f : X \to \mathbb{R}^d$. As the conditions of our results may appear a bit abstract, Remark 4.5 in Sect. 4 contains examples of systems where these conditions have been verified.

At the end of the paper the reader will find several appendices, which are integral parts of the paper: in Appendix A we interpret the limit variance σ^2 in the language of RDSs and skew products. In Appendix B we present conditions for $\sigma^2 > 0$. In Appendix C, we discuss how the fiberwise centering in the definition of $\bar{W}_n$ affects the limit variance. For completeness, in Appendix D we elaborate on the structure of an invariant measure intimately related to the problem.

2 The Term $|\sigma_n^2(\omega) - \mathbb{E}\sigma_n^2|$

In this section identify conditions which guarantee that, almost surely, $|\sigma_n^2(\omega) - \mathbb{E}\sigma_n^2|$ tends to zero at a specific rate.

Standing Assumption (*SA1*) throughout this paper we will assume that f is a bounded measurable function and μ is a probability measure. We also assume that a uniform decay of correlations holds in that

$$|\mu(\bar{f}_i \bar{f}_j)| \le \eta(|i - j|)$$

almost surely, where $\eta : \mathbb{N} \to [0, \infty)$ is such that

$$\sum_{i=0}^{\infty} \eta(i) < \infty \quad \text{and} \quad \eta \text{ is non-increasing.} \tag{2}$$

$\blacksquare$

Note already that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} i\eta(i) = 0$$

because $\lim_{i\to\infty} i\eta(i) = 0$. For the most part, we shall require additional conditions on η.

For future convenience, let us introduce the random variables

$$v_i = v_i(\omega) = \sum_{j=i}^{\infty} (2 - \delta_{ij})\mu(\bar{f}_i \bar{f}_j)$$

(where $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ otherwise) and their centered counterparts

$$\tilde{v}_i = v_i - \mathbb{E}v_i.$$

Note that these are uniformly bounded. We also denote

$$\tilde{\sigma}_n^2 = \sigma_n^2 - \mathbb{E}\sigma_n^2.$$

Thus, our objective is to show $\tilde{\sigma}_n^2 \to 0$ at some rate.

The following lemma is readily obtained by a well-known computation:

Lemma 2.1 *Assuming* (2), *there exists a constant* $C > 0$ *such that*

$$\left| \sigma_n^2 - \frac{1}{n} \sum_{i=0}^{n-1} v_i \right| \le C \left(\frac{1}{n} \sum_{i=1}^{n} i\eta(i) + \sum_{i=n+1}^{\infty} \eta(i) \right) = o(1)$$

for all ω.

Proof First, we compute

$$
\begin{aligned}
\sigma_n^2 &= \frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\mu(\bar{f}_i\bar{f}_j) = \frac{1}{n}\left[\sum_{i=0}^{n-1}\mu(\bar{f}_i^2) + 2\sum_{0\le i<j<n}\mu(\bar{f}_i\bar{f}_j)\right] \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\left[\mu(\bar{f}_i^2) + 2\sum_{j=i+1}^{n-1}\mu(\bar{f}_i\bar{f}_j)\right] \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\left[\mu(\bar{f}_i^2) + 2\sum_{j=i+1}^{\infty}\mu(\bar{f}_i\bar{f}_j)\right] + O\left(\frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=n}^{\infty}\eta(j-i)\right).
\end{aligned}
$$

Here

$$
\begin{aligned}
\frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=n}^{\infty}\eta(j-i) &= \frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=n}^{n+i}\eta(j-i) + \frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=n+i+1}^{\infty}\eta(j-i) \\
&= \frac{1}{n}\sum_{i=1}^{n} i\eta(i) + \sum_{i=n+1}^{\infty}\eta(i).
\end{aligned}
$$

The last sums tend to zero by assumption. □

Suppose that $\eta(0) = A$ and $\eta(n) = An^{-\psi}$, $n \ge 1$, for some constants $A \ge 0$, $\psi > 0$. We then use shorthand notation $\eta(n) = An^{-\psi}$, i.e., we interpret $0^{-\psi} = 1$.

Corollary 2.2 *Suppose $\eta(n) = An^{-\psi}$, where $\psi > 1$. Then*

$$
\left|\sigma_n^2 - \frac{1}{n}\sum_{i=0}^{n-1} v_i\right| \le C\begin{cases} n^{-1}, & \psi > 2, \\ n^{-1}\log n, & \psi = 2, \\ n^{1-\psi}, & 1 < \psi < 2. \end{cases}
$$

We skip the elementary proof based on Lemma 2.1.

Remark 2.3 Of course, the upper bounds in the preceding results apply equally well to

$$
\tilde{\sigma}_n^2 - \frac{1}{n}\sum_{i=0}^{n-1}\tilde{v}_i.
$$

The following result, which has been used in dynamical systems papers including Melbourne–Nicol [17], will be used to obtain an almost sure convergence rate of $\frac{1}{n}\sum_{i=0}^{n-1}\tilde{v}_i$ to zero:

Theorem 2.4 (Gál–Koksma [12]; see also Philipp–Stout [20]) *Let (X_n) be a sequence of centered, square-integrable, random variables. Suppose there exist $C > 0$ and $q > 0$ such that*

$$
\mathbb{E}\left[\left(\sum_{k=m}^{m+n-1} X_k\right)^2\right] \le C[(n+m)^q - m^q]
$$

for all $m \ge 0$ and $n \ge 1$. Let $\delta > 0$ be arbitrary. Then, almost surely,

$$
\frac{1}{n}\sum_{k=1}^{n} X_k = O\left(n^{\frac{q}{2}-1}\log^{\frac{3}{2}+\delta} n\right).
$$

Remark 2.5 In this paper the theorem is applied in the range $1 \le q < 2$. In particular, $n^q + m^q \le (n+m)^q$ then holds, so it suffices to establish an upper bound of the form Cn^q.

Our application of Theorem 2.4 will be based on the following standard lemma:

Lemma 2.6 *Suppose $|\mathbb{E}[X_i X_k]| \le r(|k-i|)$ where $r(k) = O(k^{-\beta})$. There exists a constant $C > 0$ such that*

$$\mathbb{E}\left[\left(\sum_{k=m}^{m+n-1} X_k\right)^2\right] \le C \begin{cases} n, & \beta > 1, \\ n \log n, & \beta = 1, \\ n^{2-\beta}, & 0 < \beta < 1. \end{cases}$$

Proof Note that

$$\mathbb{E}\left[\left(\sum_{k=m}^{m+n-1} X_k\right)^2\right] \le \sum_{k=m}^{m+n-1}\sum_{l=m}^{m+n-1} r(|k-l|) = nr(0) + \sum_{k=1}^{n-1} 2(n-k)r(k)$$
$$\le nr(0) + 2n\sum_{k=1}^{n-1} r(k) \le Cn\sum_{k=1}^{n-1} k^{-\beta}.$$

Bounding the last sum in each case yields the result. □

2.1 Dependent Random Selection Process

It is most interesting to study the case where the sequence $\omega = (\omega_i)_{i\ge 1}$ is generated by a nontrivial stochastic process such that the measure $\mathbb{P}$ is not the product of its one-dimensional marginals. Essentially without loss of generality, we pass directly to the so-called canonical version of the process, which corresponds to the point of view that the sequence ω *is* the seed of the random process. In the following we briefly review some standard details.

Let $\pi_i : \Omega \to \Omega_0$ be the projection $\pi_i(\omega) = \omega_i$. The product sigma-algebra $\mathcal{F}$ is the smallest sigma-algebra with respect to which all the latter projections are measurable. For any $I = (i_1, \dots, i_p) \subset \mathbb{Z}_+$, $p \in \mathbb{Z}_+ \cup \{\infty\}$, we may define the sub-sigma-algebra $\mathcal{F}_I = \sigma(\pi_i : i \in I)$ of $\mathcal{F}$. (In particular, $\mathcal{F} = \mathcal{F}_{\mathbb{Z}_+}$.) We also recall that a function $u : \Omega \to \mathbb{R}$ is $\mathcal{F}_I$-measurable if and only if there exists an $\mathcal{E}^p$-measurable function $\tilde{u} : \Omega_0^p \to \mathbb{R}$ such that $u = \tilde{u} \circ (\pi_{i_1}, \dots, \pi_{i_p})$, i.e., $u(\omega) = \tilde{u}(\omega_{i_1}, \dots, \omega_{i_p})$. With slight abuse of language, we will say below that the sigma-algebra $\mathcal{F}_I$ is generated by the random variables ω_i, $i \in I$, instead of the projections π_i. In particular, we denote

$$\mathcal{F}_i^j = \sigma(\omega_n : i \le n \le j) \subset \mathcal{F}$$

for $1 \le i \le j \le \infty$.

Denote

$$\alpha\left(\mathcal{F}_1^i, \mathcal{F}_j^\infty\right) = \sup_{A \in \mathcal{F}_1^i,\, B \in \mathcal{F}_j^\infty} |\mathbb{P}(AB) - \mathbb{P}(A)\,\mathbb{P}(B)|.$$

In the following $(\alpha(n))_{n\ge 1}$ will denote a sequence such that

$$\sup_{i\ge 1} \alpha\left(\mathcal{F}_1^i, \mathcal{F}_{i+n}^\infty\right) \le \alpha(n)$$

for each $n \ge 1$.

Standing Assumption (*SA2*) throughout the rest of the paper we assume that the random selection process is **strong mixing**: $\alpha(n)$ can be chosen so that[1]

$$\lim_{n\to\infty} \alpha(n) = 0 \quad \text{and} \quad \alpha \text{ is non-increasing.}$$

■

Suppose that $u = u(\omega_1, \dots, \omega_i)$ and $v = v(\omega_{i+n}, \omega_{i+n+1}, \dots)$ are L^∞ functions. Then

$$|\mathbb{E}[uv] - \mathbb{E}u\mathbb{E}v| \le 4\|u\|_\infty \|v\|_\infty \alpha(n) \tag{3}$$

as is well known. Ultimately, we will impose a rate of decay on $\alpha(n)$.

We denote by T_* the pushforward of a map T, acting on a probability measure m, i.e., $(T_*m)(A) = m(T^{-1}A)$ for measurable sets A. We write

$$\mu_k = \left(T_{\omega_k} \circ \dots \circ T_{\omega_1}\right)_* \mu$$

and

$$\mu_{k,r+1} = \left(T_{\omega_k} \circ \dots \circ T_{\omega_{r+1}}\right)_* \mu$$

for $k \ge r$. We also write

$$f_{l,k+1} = f \circ T_{\omega_l} \circ \dots \circ T_{\omega_{k+1}} = f \circ \varphi(l-k, \tau^k \omega)$$

for $l \ge k$. Note that all of these objects depend on ω through the maps T_{ω_i}. We use the conventions $\mu_0 = \mu$, $\mu_{r,r+1} = \mu$ and $f_{k,k+1} = f$ here.
Standing Assumption (*SA3*) throughout the rest of the paper we assume the following uniform **memory-loss condition**: there exists a constant $C \ge 0$ such that

$$|\mu_k(g) - \mu_{k,r+1}(g)| \le C\eta(k-r) \tag{4}$$

for all

$$g \in \mathcal{G}_k = \mathcal{G}_k(\omega) = \{f_{l,k+1} : l \ge k\} \cup \{f f_{l,k+1} : l \ge k\}$$

whenever $k \ge r$. The bound holds uniformly for (almost) all ω. ■

In the cocycle notation, (4) reads

$$|\mu(g \circ \varphi(k, \omega)) - \mu(g \circ \varphi(k-r, \tau^r \omega))| \le C\eta(k-r). \tag{5}$$

Note that, setting

$$\tilde{c}_{ij} = (2 - \delta_{ij})[\mu(\bar{f}_i \bar{f}_j) - \mathbb{E}\mu(\bar{f}_i \bar{f}_j)],$$

we have

$$\tilde{v}_i = \sum_{j=i}^{\infty} \tilde{c}_{ij} \quad \text{and} \quad \mathbb{E}[\tilde{v}_i \tilde{v}_k] = \mathbb{E}\left[\left(\sum_{j=i}^{\infty} \tilde{c}_{ij}\right)\left(\sum_{l=k}^{\infty} \tilde{c}_{kl}\right)\right].$$

Lemma 2.7 *There exists a constant $C \ge 0$ such that*

$$|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \le \begin{cases} C\eta(j-i)\eta(l-k), & \text{if } i \le j \text{ and } k \le l, \\ C\eta(j-i)\min_{r: j \le r \le k}\{\eta(k-r) + \alpha(r-j)\eta(l-k)\}, & \text{if } i \le j \le k \le l. \end{cases}$$

[1] It would be standard to denote $\sup_{i\ge 1} \alpha(\mathcal{F}_1^i, \mathcal{F}_{i+n}^\infty)$ by $\alpha(n)$. We prefer to let $\alpha(n)$ stand for an upper bound so the non-increasing assumption makes sense. This is a choice of technical convenience.

In particular, for $i \le j \le k \le l$,

$$|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \le C\eta(j-i)\min\left\{\eta(l-k),\ \min_{r:j\le r\le k}\{\eta(k-r)+\alpha(r-j)\eta(l-k)\}\right\}.$$

Proof The first bound holds, because

$$|\tilde{c}_{ij}\tilde{c}_{kl}| \le 2\eta(|j-i|) \quad \text{and} \quad |\tilde{c}_{ij}\tilde{c}_{kl}| \le 2\eta(|l-k|).$$

Suppose $i \le j \le r \le k \le l$ holds. By (SA3), the choice $g = ff_{l,k+1}$ yields

$$|\mu(f_k f_l) - \mu_{k,r+1}(ff_{l,k+1})| \le C\eta(k-r),$$

while the choices $g = f$ and $g = f_{l,k+1}$ together yield

$$|\mu(f_k)\mu(f_l) - \mu_{k,r+1}(f)\mu_{k,r+1}(f_{l,k+1})| \le C\eta(k-r).$$

Hence

$$|\mu(\bar{f}_k\bar{f}_l) - \{\mu_{k,r+1}(ff_{l,k+1}) - \mu_{k,r+1}(f)\mu_{k,r+1}(f_{l,k+1})\}| \le C\eta(k-r). \tag{6}$$

Note that here the expression in the curly braces only depends on the random variables $\omega_{r+1},\dots,\omega_l$ while $\mu(\bar{f}_i\bar{f}_j)$ only depends on $\omega_1,\dots,\omega_j$. More precisely, denoting $u = \mu(\bar{f}_i\bar{f}_j)$ and $v = \mu_{k,r+1}(ff_{l,k+1}) - \mu_{k,r+1}(f)\mu_{k,r+1}(f_{l,k+1})$, we have $u \in L^\infty(\mathcal{F}_1^j)$ and $v \in L^\infty(\mathcal{F}_{r+1}^l) \subset L^\infty(\mathcal{F}_r^\infty)$. Therefore,

$$|\mathbb{E}[\mu(\bar{f}_i\bar{f}_j)\mu(\bar{f}_k\bar{f}_l)] - \mathbb{E}[uv]| \le C\mathbb{E}[|u|]\eta(k-r) \le C\eta(j-i)\eta(k-r)$$

by (6). On the other hand, the strong-mixing bound (3) implies

$$|\mathbb{E}[uv] - \mathbb{E}u\,\mathbb{E}v| \le \alpha(r-j)\|u\|_\infty\|v\|_\infty \le C\alpha(r-j)\eta(j-i)\|v\|_\infty.$$

Moreover,

$$|\mathbb{E}[\mu(\bar{f}_i\bar{f}_j)]\mathbb{E}[\mu(\bar{f}_k\bar{f}_l)] - \mathbb{E}u\,\mathbb{E}v| \le |\mathbb{E}u||\mathbb{E}[\mu(\bar{f}_k\bar{f}_l) - v]| \le C\eta(j-i)\eta(k-r).$$

Collecting the bounds leads to the estimate

$$\begin{aligned}|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| &\le 4|\mathbb{E}[\mu(\bar{f}_i\bar{f}_j)\mu(\bar{f}_k\bar{f}_l)] - \mathbb{E}[\mu(\bar{f}_i\bar{f}_j)]\mathbb{E}[\mu(\bar{f}_k\bar{f}_l)]| \\ &\le C\eta(j-i)\{\eta(k-r)+\alpha(r-j)\|v\|_\infty\}.\end{aligned}$$

Note that (6) immediately yields the estimate

$$\|v\|_\infty \le C\eta(l-k) + C\eta(k-r)$$

which by the boundedness of α results in

$$\begin{aligned}|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| &\le C\eta(j-i)\{\eta(k-r)+\alpha(r-j)[\eta(l-k)+\eta(k-r)]\} \\ &\le C\eta(j-i)\{\eta(k-r)+\alpha(r-j)\eta(l-k)\}.\end{aligned}$$

Taking the minimum with respect to r proves the lemma. □

The upper bound $|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \le C\eta(j-i)\eta(l-k)$ of Lemma 2.7 yields the following intermediate result:

Lemma 2.8 *For $i \le k$,*

$$|\mathbb{E}[\tilde{v}_i \tilde{v}_k]| \le C \left(\sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| + \sum_{n=m}^{\infty} \eta(n) + \sum_{n=0}^{m-1} \sum_{p=m-n}^{\infty} \eta(n)\eta(p) \right),$$

where we have denoted $m = k - i$.

Proof We can estimate

$$\begin{aligned}
|\mathbb{E}[\tilde{v}_i \tilde{v}_k]| &\le \sum_{j=i}^{\infty} \sum_{l=k}^{\infty} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \\
&= \sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| + \sum_{j=i}^{k-1} \sum_{l=2k-j}^{\infty} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| + \sum_{j=k}^{\infty} \sum_{l=k}^{\infty} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \\
&\le \sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| + \sum_{j=i}^{k-1} \sum_{l=2k-j}^{\infty} C\eta(j-i)\eta(l-k) + \sum_{j=k}^{\infty} \sum_{l=k}^{\infty} C\eta(j-i)\eta(l-k) \\
&\le \sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| + C \sum_{n=0}^{k-i-1} \sum_{p=k-i-n}^{\infty} \eta(n)\eta(p) + C \sum_{n=k-i}^{\infty} \eta(n).
\end{aligned}$$

In the third line we used the upper bound $|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \le C\eta(j-i)\eta(l-k)$ of Lemma 2.7. ◻

Next we investigate the remaining term $\sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]|$ appearing in Lemma 2.8. Since $i \le j \le k \le l$, we have

$$\min_{r:j\le r\le k} \{\eta(k-r) + \alpha(r-j)\eta(l-k)\} \le \eta(k-j) + \alpha(0)\eta(l-k)$$

by choosing $r = j$. Suppose furthermore that $k - j \ge l - k$ and recall η is non-increasing. Then the right side of the above display is bounded above by $C\eta(l-k)$. In other words, if $i \le j \le k \le l \le 2k - j$, then $C\eta(j-i)\min_{r:j\le r\le k}\{\eta(k-r) + \alpha(r-j)\eta(l-k)\}$ is the tightest bound on $|\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]|$ that Lemma 2.7 can provide. This observation motivates the following lemma.

Lemma 2.9 *Define*

$$S(i,k) = \sum_{j=i}^{k-1} \eta(j-i) \sum_{l=k}^{2k-j-1} \min_{r:j\le r\le k} \{\eta(k-r) + \alpha(r-j)\eta(l-k)\}.$$

(i) *There exists a constant $C \ge 0$ such that*

$$\sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \le CS(i,k)$$

whenever $i \le k$.
(ii) *There exist constants $C_1 \ge 0$ and $C_2 \ge 0$ such that*

$$C_1\{m\eta(m) + \alpha(m)\} \le S(i,k) \le C_2\left\{m\eta\left(\left\lfloor \frac{m}{4} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m}{4} \right\rfloor\right)\right\} \quad (m = k - i)$$

whenever $i < k$. (Note also that $S(i,i) = 0$).

Proof Part (i) is an immediate corollary of Lemma 2.7. As for part (ii), let us first prove the lower bound. Since all the terms in $S(i,k)$ are nonnegative and α is non-increasing, we have for $i<k$ that

$$S(i,k) \ge \sum_{j=i}^{i} \eta(j-i) \sum_{l=k}^{k} \min_{r:j\le r\le k} \alpha(r-j)\eta(l-k) \ge \eta(0)^2\alpha(k-i)$$

and

$$S(i,k) \ge \sum_{j=i}^{i} \eta(j-i) \sum_{l=k}^{2k-j-1} \min_{r:j\le r\le k} \eta(k-r) \ge \eta(0)(k-i)\eta(k-i).$$

Setting $C_1 = \frac{1}{2}\eta(0)^2 + \frac{1}{2}\eta(0)$ gives an overall bound

$$S(i,k) \ge C_1\{\alpha(m) + m\eta(m)\}$$

for all $i<k$.

It remains to prove the upper bound in part (ii). We choose $r = \lfloor (k+j)/2 \rfloor$. Since η is summable, we have

$$\begin{aligned}
S(i,k) &= \sum_{j=i}^{k-1} \eta(j-i) \sum_{l=k}^{2k-j-1} \left\{ \eta\left(k - \left\lfloor \frac{k+j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{k+j}{2} \right\rfloor - j\right)\eta(l-k) \right\} \\
&\le C \sum_{j=i}^{k-1} \eta(j-i) \left\{ (k-j)\eta\left(\left\lfloor \frac{k-j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{k-j}{2} \right\rfloor\right) \right\} \\
&= C \sum_{j=0}^{m-1} \eta(j) \left\{ (m-j)\eta\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) \right\}.
\end{aligned}$$

Next we split the last sum above into two parts, keeping in mind that α and η are non-increasing and η is also summable:

$$\begin{aligned}
& C \sum_{j=0}^{m-1} \eta(j) \left\{ (m-j)\eta\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) \right\} \\
&\le C \sum_{j=0}^{\lfloor m/2 \rfloor} \eta(j) \left\{ (m-j)\eta\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) \right\} \\
&\quad + C \sum_{j=\lfloor m/2 \rfloor+1}^{m-1} \eta(j) \left\{ (m-j)\eta\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) \right\} \\
&\le C \sum_{j=0}^{\lfloor m/2 \rfloor} \eta(j) \left\{ m\eta\left(\left\lfloor \frac{m}{4} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m}{4} \right\rfloor\right) \right\} \\
&\quad + C\eta\left(\left\lfloor \frac{m}{2} \right\rfloor\right) \sum_{j=m/2}^{m} \left\{ m\eta\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m-j}{2} \right\rfloor\right) \right\} \\
&\le C \left\{ m\eta\left(\left\lfloor \frac{m}{4} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m}{4} \right\rfloor\right) \right\} + Cm\eta\left(\left\lfloor \frac{m}{2} \right\rfloor\right) \le C_2 \left\{ m\eta\left(\left\lfloor \frac{m}{4} \right\rfloor\right) + \alpha\left(\left\lfloor \frac{m}{4} \right\rfloor\right) \right\}.
\end{aligned}$$

This completes the proof. □

The next two lemmas concern the case when η and α are polynomial.

Lemma 2.10 *Let* $\eta(n) = Cn^{-\psi}$, $\psi > 1$ *and* $\alpha(n) = Cn^{-\gamma}$, $\gamma > 0$. *Then*

$$C_1 m^{-\min\{\psi-1,\gamma\}} \le S(i,k) \le C_2 m^{-\min\{\psi-1,\gamma\}} \quad (m = k - i)$$

whenever $i < k$.

Proof The lower bound follows immediately from Lemma 2.9(ii). Let first $m \ge 8$. Then $\lfloor m/4 \rfloor \ge m/8$. Thus Lemma 2.9(ii) yields

$$S(i,k) \le C\left\{m\eta\left(\frac{m}{8}\right) + \alpha\left(\frac{m}{8}\right)\right\} \le Cm^{\max\{1-\psi,-\gamma\}},$$

when $m \ge 8$. Since $S(i,k) \le C(k-i)^2 = Cm^2$ by counting terms, we can choose a large enough C_2 such that the claimed upper bound holds also for $1 \le m < 8$. □

Lemma 2.11 *Let* $\eta(n) = Cn^{-\psi}$, $\psi > 1$ *and* $\alpha(n) = Cn^{-\gamma}$, $\gamma > 0$. *Then*

$$|\mathbb{E}[\tilde{v}_i \tilde{v}_k]| \le Cm^{-\min\{\psi-1,\gamma\}} \quad (m = k - i)$$

whenever $i < k$.

Proof Firstly,

$$\sum_{n=m}^{\infty} \eta(n) \le Cm^{1-\psi}. \tag{7}$$

Secondly,

$$\begin{aligned}
\sum_{n=0}^{m-1} \sum_{p=m-n}^{\infty} \eta(n)\eta(p) &= \sum_{p=m}^{\infty} \eta(p) + \sum_{n=1}^{m-1} \sum_{p=m-n}^{\infty} \eta(n)\eta(p) \\
&= C\sum_{p=m}^{\infty} p^{-\psi} + C\sum_{n=1}^{m-1} \sum_{p=m-n}^{\infty} n^{-\psi} p^{-\psi} \\
&\le Cm^{1-\psi} + C\sum_{n=1}^{m-1} n^{-\psi}(m-n)^{1-\psi} \\
&= Cm^{1-\psi} + C\sum_{n=1}^{m-1} n^{1-\psi}(m-n)^{-\psi}.
\end{aligned}$$

Regarding the last sum appearing above, observe that

$$\sum_{n=1}^{m/2} n^{1-\psi}(m-n)^{-\psi} \le \sum_{n=1}^{m/2} 1^{1-\psi}(m/2)^{-\psi} \le Cm^{1-\psi},$$

while

$$\sum_{n=m/2}^{m-1} n^{1-\psi}(m-n)^{-\psi} \le \sum_{n=m/2}^{m-1} (m/2)^{1-\psi}(m-n)^{-\psi} \le (m/2)^{1-\psi} \sum_{n=1}^{m/2} n^{-\psi} \le Cm^{1-\psi}.$$

In other words, also

$$\sum_{n=0}^{m-1} \sum_{p=m-n}^{\infty} \eta(n)\eta(p) \le C m^{1-\psi}. \tag{8}$$

Now, by Lemmas 2.9(i) and 2.10 we have

$$\sum_{j=i}^{k-1} \sum_{l=k}^{2k-j-1} |\mathbb{E}[\tilde{c}_{ij}\tilde{c}_{kl}]| \le S(i,k) \le C m^{\max\{1-\psi,-\gamma\}}.$$

Thus, Lemma 2.8 and bounds (7) and (8) yield

$$|\mathbb{E}[\tilde{v}_i \tilde{v}_k]| \le C m^{1-\psi} + C m^{\max\{1-\psi,-\gamma\}} \le C m^{\max\{1-\psi,-\gamma\}}.$$

The proof is complete. □

Lemma 2.12 *Suppose* $|\mathbb{E}[\tilde{v}_i \tilde{v}_k]| \le C(k-i)^{-\beta}$ *for all* $i < k$. *Let* $\delta > 0$ *be arbitrary. Then*

$$\frac{1}{n}\sum_{k=1}^{n} \tilde{v}_k = \begin{cases} O\left(n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n\right), & \beta > 1, \\ O\left(n^{-\frac{1}{2}+\delta}\right), & \beta = 1, \\ O\left(n^{-\frac{\beta}{2}} \log^{\frac{3}{2}+\delta} n\right), & 0 < \beta < 1 \end{cases}$$

almost surely.

Proof Applying Lemma 2.6 we get

$$\mathbb{E}\left[\left(\sum_{k=m}^{m+n-1} \tilde{v}_k\right)^2\right] \le C \begin{cases} n, & \beta > 1, \\ n\log n, & \beta = 1, \\ n^{2-\beta}, & 0 < \beta < 1. \end{cases}$$

Notice that for any $\varepsilon > 0$ we have $n \log n = O(n^{1+\varepsilon})$. Applying Theorem 2.4 with

$$q = \begin{cases} 1, & \beta > 1, \\ 1+\delta, & \beta = 1, \\ 2-\beta, & 0 < \beta < 1, \end{cases}$$

yields the claim. □

Proposition 2.13 *Let* $\eta(n) = Cn^{-\psi}$, $\psi > 1$ *and* $\alpha(n) = Cn^{-\gamma}$, $\gamma > 0$. *Then, for any* $\delta > 0$,

$$\frac{1}{n}\sum_{k=1}^{n} \tilde{v}_k = \begin{cases} O\left(n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n\right), & \min\{\psi-1,\gamma\} > 1, \\ O\left(n^{-\frac{1}{2}+\delta}\right), & \min\{\psi-1,\gamma\} = 1, \\ O\left(n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n\right), & 0 < \min\{\psi-1,\gamma\} < 1, \end{cases}$$

almost surely.

Proof By Lemma 2.11, we have $|\mathbb{E}[\tilde{v}_i \tilde{v}_k]| \le C m^{-\min\{\psi-1,\gamma\}}$. Applying Lemma 2.12 with $\beta = \min\{\psi-1,\gamma\}$ yields the claim. □

We are now in position to prove the main result of this section:

Theorem 2.14 *Assume* (SA1) *and* (SA3) *with* $\eta(n) = Cn^{-\psi}$, $\psi > 1$. *Assume* (SA2) *with* $\alpha(n) = Cn^{-\gamma}$, $\gamma > 0$. *Then, for arbitrary* $\delta > 0$,

$$\left|\sigma_n^2 - \mathbb{E}\sigma_n^2\right| = \begin{cases} O\left(n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n\right), & \min\{\psi-1,\gamma\} > 1, \\ O\left(n^{-\frac{1}{2}+\delta}\right), & \min\{\psi-1,\gamma\} = 1, \\ O\left(n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n\right), & 0 < \min\{\psi-1,\gamma\} < 1, \end{cases}$$

almost surely.

Proof By Corollary 2.2,

$$\left|\sigma_n^2 - \mathbb{E}\sigma_n^2 - \frac{1}{n}\sum_{i=0}^{n-1} \tilde{v}_i\right| \le C \begin{cases} n^{-1}, & \psi > 2, \\ n^{-1}\log n, & \psi = 2, \\ n^{1-\psi}, & 1 < \psi < 2. \end{cases}$$

Combining this with Proposition 2.13 yields the following upper bounds on $|\sigma_n^2 - \mathbb{E}\sigma_n^2|$:

$$\begin{cases} O\left(n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n + n^{-1}\right), & \min\{\psi-1,\gamma\} > 1, \\ O\left(n^{-\frac{1}{2}+\delta} + n^{-1}\log n\right), & \min\{\psi-1,\gamma\} = 1, \\ O\left(n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n + n^{-\min\{1,\psi-1\}} + n^{-1}\log n\right), & 0 < \min\{\psi-1,\gamma\} < 1. \end{cases}$$

In each case the first term is the largest, so the proof is complete. □

3 The Term $|\mathbb{E}\sigma_n^2 - \sigma^2|$

In this section we formulate general condition that allow to identify the limit $\sigma^2 = \lim_{n\to\infty} \mathbb{E}\sigma_n^2$ and obtain a rate of convergence.

Write

$$c_{ij} = \mu(f_i f_j) - \mu(f_i)\mu(f_j)$$

for brevity. Then

$$\sigma_n^2 = \frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} c_{ij} = \frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=i}^{n-1}(2-\delta_{ij})c_{ij} = \frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{n-1-i}(2-\delta_{i,i+k})c_{i,i+k}.$$

Setting[2]

$$v_{ik} = (2-\delta_{k0})[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})]$$

we arrive at

$$\sigma_n^2 = \frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{n-1-i} v_{ik} \quad \text{and} \quad \mathbb{E}\sigma_n^2 = \frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{n-1-i} \mathbb{E}v_{ik}.$$

Recall that

$$|v_{ik}| \le 2\eta(k). \tag{9}$$

[2] The relationship with earlier notations is that $\sum_{k=0}^{\infty} v_{ik} = v_i$.

3.1 Asymptotics of Related Double Sums of Real Numbers

In this subsection we consider double sequences of uniformly bounded numbers a_{ik}, $(i, k) \in \mathbb{N}^2$, with the objective of controlling the sequence

$$B_n = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{k=0}^{n-1-i} a_{ik}$$

for large values of n. In this subsection, we make the following assumption tailored to our later needs:
Standing assumption there exists $\eta : \mathbb{N} \to [0, \infty)$ such that

$$|a_{ik}| \le \eta(k) \quad \text{and} \quad \sum_{k=0}^{\infty} \eta(k) < \infty. \tag{10}$$

We also denote the tail sums of η by

$$R(K) = \sum_{k=K+1}^{\infty} \eta(k).$$

We begin with a handy observation:

Lemma 3.1 *There exists $C > 0$ such that*

$$\left| B_n - \frac{1}{n} \sum_{i=0}^{n-1} \sum_{k=0}^{L} a_{ik} \right| \le C(R(K) + Kn^{-1})$$

whenever $0 < K \le n$ *and* $K \le L \le \infty$.

Proof For all choices of $0 < K \le n$ we have

$$\begin{aligned}
B_n &= \frac{1}{n} \sum_{i=0}^{n-K-1} \sum_{k=0}^{n-1-i} a_{ik} + \frac{1}{n} \sum_{i=n-K}^{n-1} \sum_{k=0}^{n-1-i} a_{ik} = \frac{1}{n} \sum_{i=0}^{n-K-1} \sum_{k=0}^{n-1-i} a_{ik} + O(Kn^{-1}) \\
&= \frac{1}{n} \sum_{i=0}^{n-K-1} \sum_{k=0}^{K} a_{ik} + \frac{1}{n} \sum_{i=0}^{n-K-1} \sum_{k=K+1}^{n-1-i} a_{ik} + O(Kn^{-1}) \\
&= \frac{1}{n} \sum_{i=0}^{n-K-1} \sum_{k=0}^{K} a_{ik} + O(R(K) + Kn^{-1}) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \sum_{k=0}^{K} a_{ik} - \frac{1}{n} \sum_{i=n-K}^{n-1} \sum_{k=0}^{K} a_{ik} + O(R(K) + Kn^{-1}) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \sum_{k=0}^{K} a_{ik} + O(R(K) + Kn^{-1}).
\end{aligned}$$

The error is uniform because of the uniform condition $|a_{ik}| \le \eta(k)$. For $L \ge K$,

$$\sum_{k=0}^{L} a_{ik} - \sum_{k=0}^{K} a_{ik} = \sum_{k=K+1}^{L} a_{ik} = O(R(K))$$

uniformly, which concludes the proof. □

The following lemma helps identify the limit of B_n and the rate of convergence under certain circumstances:

Lemma 3.2 *Suppose the limit*

$$b_k = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} a_{ik}$$

exists for all $k \geq 0$. Then

$$\lim_{n\to\infty} B_n = \sum_{k=0}^{\infty} b_k.$$

The series on the right side converges absolutely. Furthermore, denoting

$$r_k(n) = \frac{1}{n} \sum_{i=0}^{n-1} a_{ik} - b_k$$

there exists $C > 0$ such that

$$\left| B_n - \sum_{k=0}^{\infty} b_k \right| \leq C \left(\left| \sum_{k=0}^{K} r_k(n) \right| + R(K) + Kn^{-1} \right) \tag{11}$$

holds whenever $0 < K \leq n$.

Proof Since

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} a_{ik} \right| \leq \eta(k),$$

also $|b_k| \leq \eta(k)$, so the series $\sum_{k=0}^{\infty} b_k$ converges absolutely. Lemma 3.1 with $L = K$ yields

$$B_n = \sum_{k=0}^{K} \frac{1}{n} \sum_{i=0}^{n-1} a_{ik} + O(R(K) + Kn^{-1})$$

uniformly for all $0 < K \leq n$. Thus, the definition of $r_k(n)$ gives

$$B_n = \sum_{k=0}^{K} b_k + O\left(\left| \sum_{k=0}^{K} r_k(n) \right| + R(K) + Kn^{-1} \right).$$

Now $|b_k| \leq \eta(k)$ yields (11). To prove the convergence of B_n, consider (11) and fix an arbitrary $\varepsilon > 0$. Fix K so large that $R(K) < \varepsilon/2C$. Since $\left|\sum_{k=0}^{K} r_k(n)\right| + Kn^{-1}$ tends to zero with increasing n, it is bounded by $\varepsilon/2C$ for all large n. Then $|B_n - \sum_{k=0}^{\infty} b_k| < \varepsilon$. □

3.2 Convergence of $\mathbb{E}\sigma_n^2$: A General Result

In this subsection we apply the results of the preceding subsection to the sequence

$$\mathbb{E}\sigma_n^2 = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{k=0}^{n-1-i} \mathbb{E}v_{ik}$$

where

$$\mathbb{E}v_{ik} = (2 - \delta_{k0})\,\mathbb{E}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})].$$

Recall from (9) and (2) of (SA1) that the standing assumption in (10) is satisfied: $|\mathbb{E}v_{ik}| \le 2\eta(k)$ and $\sum_{k=0}^{\infty}\eta(k) < \infty$. The next theorem is nothing but a rephrasing of Lemma 3.2 in the case $a_{ik} = \mathbb{E}v_{ik}$ at hand.

Theorem 3.3 *Suppose the limit*

$$V_k = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}v_{ik}$$

exists for all $k \ge 0$. *The series*

$$\sigma^2 = \sum_{k=0}^{\infty} V_k$$

is absolutely convergent, and

$$\lim_{n\to\infty}\mathbb{E}\sigma_n^2 = \sigma^2.$$

In particular, $\sigma^2 \ge 0$. *Furthermore, there exists a constant* $C > 0$ *such that*

$$\left|\mathbb{E}\sigma_n^2 - \sigma^2\right| \le C\left(\left|\sum_{k=0}^{K}\left(\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}v_{ik} - V_k\right)\right| + \sum_{k=K+1}^{\infty}\eta(k) + Kn^{-1}\right)$$

holds whenever $0 < K \le n$.

3.3 Convergence of $\mathbb{E}\sigma_n^2$: Asymptotically Mean Stationary $\mathbb{P}$

For the rest of the section we assume $\mathbb{P}$ is asymptotically mean stationary, with mean $\bar{\mathbb{P}}$. In other words, there exists a measure $\bar{\mathbb{P}}$ such that, given a bounded measurable $g : \Omega \to \mathbb{R}$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\int g\circ\tau^i\,\mathrm{d}\mathbb{P} = \int g\,\mathrm{d}\bar{\mathbb{P}}. \tag{12}$$

The measure $\bar{\mathbb{P}}$ is then τ-invariant. We denote $\bar{\mathbb{E}}g = \int g\,\mathrm{d}\bar{\mathbb{P}}$. We will shortly impose additional rate conditions; see (15).

Recall the cocycle property of the random compositions. In what follows, it will be convenient to use the notations

$$g_{ik}^1(\omega) = \mu(f_i f_{i+k}) = \mu(f\circ\varphi(i,\omega)\,f\circ\varphi(i+k,\omega))$$

and

$$g_{ik}^2(\omega) = \mu(f_i)\mu(f_{i+k}) = \mu(f\circ\varphi(i,\omega))\mu(f\circ\varphi(i+k,\omega)).$$

For the results of this section we need the following preliminary lemma, which crucially relies on the memory-loss property (SA3), assumed to hold throughout this text.

Lemma 3.4 *There exists a constant $C > 0$ such that*

$$\left|g_{ik}^{a} - g_{rk}^{a} \circ \tau^{i-r}\right| \le C\eta(r) \tag{13}$$

for all $0 \le r \le i,\ k \ge 0$ and $a \in \{1, 2\}$.

Proof Note that we may rewrite the memory-loss property in (5) as

$$|\mu(g \circ \varphi(j, \omega)) - \mu(g \circ \varphi(r, \tau^{j-r}\omega))| \le C\eta(r),$$

for all $r \le j$. Thus, choosing $g = f$ (recall $f = f_{j,j+1} \in \mathcal{G}_j$ for all j) yields

$$\begin{aligned}
&\left|g_{ik}^{2} - g_{rk}^{2} \circ \tau^{i-r}\right| \\
&\quad = |\mu(f \circ \varphi(i, \omega))\mu(f \circ \varphi(i+k, \omega)) - \mu(f \circ \varphi(r, \tau^{i-r}\omega))\mu(f \circ \varphi(r+k, \tau^{i-r}\omega))| \\
&\quad \le |\mu(f \circ \varphi(i, \omega))||\mu(f \circ \varphi(i+k, \omega)) - \mu(f \circ \varphi(r+k, \tau^{i-r}\omega))| \\
&\qquad + |\mu(f \circ \varphi(i, \omega)) - \mu(f \circ \varphi(r, \tau^{i-r}\omega))||\mu(f \circ \varphi(r+k, \tau^{i-r}\omega))| \\
&\quad \le C(\eta(r+k) + \eta(r)) \le C\eta(r).
\end{aligned}$$

On the other hand, choosing $g = ff_{i+k,i+1} = ff \circ \varphi(k, \tau^{i}\omega)$ gives

$$\begin{aligned}
&\left|g_{ik}^{1} - g_{rk}^{1} \circ \tau^{i-r}\right| \\
&\quad = |\mu(f \circ \varphi(i, \omega) f \circ \varphi(i+k, \omega)) - \mu(f \circ \varphi(r, \tau^{i-r}\omega) f \circ \varphi(r+k, \tau^{i-r}\omega))| \\
&\quad = |\mu(g \circ \varphi(i, \omega)) - \mu(g \circ \varphi(r, \tau^{i-r}\omega))| \le C\eta(r),
\end{aligned}$$

which completes the proof. □

The following lemma guarantees that both limits $\lim_{n\to\infty} n^{-1} \sum_{i=0}^{n-1} \mathbb{E}\mu(f_i f_{i+k})$ and $\lim_{n\to\infty} n^{-1} \sum_{i=0}^{n-1} \mathbb{E}\mu(f_i)\mu(f_{i+k})$ exist and can be expressed in terms of $\bar{\mathbb{P}}$.

Lemma 3.5 *For all $i, k \ge 0$ and $a \in \{1, 2\}$*

$$\lim_{n\to\infty} n^{-1} \sum_{i=0}^{n-1} \mathbb{E}g_{ik}^{a} = \lim_{j\to\infty} \bar{\mathbb{E}}g_{jk}^{a}.$$

In particular, the limits exist.

Proof First we make the observation that since $\bar{\mathbb{P}}$ is stationary, (13) implies

$$\left|\bar{\mathbb{E}}g_{ik}^{a} - \bar{\mathbb{E}}g_{rk}^{a}\right| \le C\eta(r)$$

whenever $i \ge r$. From assumption (2) it follows that $\lim_{r\to\infty} \eta(r) = 0$. The sequence $(\bar{\mathbb{E}}g_{ik}^{a})_{i=0}^{\infty}$ is therefore Cauchy, so $\lim_{i\to\infty} \bar{\mathbb{E}}g_{ik}^{a}$ exists and respects the same bound, i.e.,

$$\left|\bar{\mathbb{E}}g_{rk}^{a} - \lim_{i\to\infty} \bar{\mathbb{E}}g_{ik}^{a}\right| \le C\eta(r). \tag{14}$$

We are now ready to show that $\lim_{n\to\infty} n^{-1} \sum_{i=0}^{n-1} \mathbb{E}g_{ik}^{a}$ exists and in the process we see that it is equal to $\lim_{j\to\infty} \bar{\mathbb{E}}g_{jk}^{a}$.

Let $\varepsilon > 0$. Choose $r \in \mathbb{N}$ such that $C\eta(r) < \varepsilon/5$, where C is the same constant as above. Then choose $n_0 \in \mathbb{N}$ that satisfies two following conditions. First, $\|f\|_\infty^2 r/n_0 < \varepsilon/5$.

Second, by (12), $\left|n^{-1}\sum_{i=0}^{n-1}\mathbb{E}g_{rk}^a\circ\tau^i-\bar{\mathbb{E}}g_{rk}^a\right|<\varepsilon/5$ for all $n\geq n_0$. Next we show that $\left|n^{-1}\sum_{i=0}^{n-1}\mathbb{E}g_{ik}^a-\lim_{j\to\infty}\bar{\mathbb{E}}g_{ik}^a\right|<\varepsilon$ for all $n\geq n_0$.

The following five estimates yield the desired result:

In this first estimate, note that $\|g_{ik}^a\|_\infty\leq\|f\|_\infty^2$ for all $i,k\in\mathbb{N}$ and $a\in\{1,2\}$:

$$\left|\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}g_{ik}^a-\frac{1}{n}\sum_{i=r}^{n-1}\mathbb{E}g_{ik}^a\right|\leq\|f^2\|_\infty n^{-1}r<\frac{\varepsilon}{5}.$$

In the second estimate, we apply (13):

$$\left|\frac{1}{n}\sum_{i=r}^{n-1}\mathbb{E}g_{ik}^a-\frac{1}{n}\sum_{i=0}^{n-r-1}\mathbb{E}g_{rk}^a\circ\tau^i\right|=\left|\frac{1}{n}\sum_{i=r}^{n-1}\mathbb{E}g_{ik}^a-\frac{1}{n}\sum_{i=r}^{n-1}\mathbb{E}g_{rk}^a\circ\tau^{i-r}\right|\leq C\eta(r)<\frac{\varepsilon}{5}.$$

The third estimate follows the same reasoning as the first:

$$\left|\frac{1}{n}\sum_{i=0}^{n-r-1}\mathbb{E}g_{rk}^a\circ\tau^i-\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}g_{rk}^a\circ\tau^i\right|\leq\|f\|_\infty^2 n^{-1}r<\frac{\varepsilon}{5}.$$

The fourth estimate follows by the definition of n_0:

$$\left|\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}g_{rk}^a\circ\tau^i-\bar{\mathbb{E}}g_{rk}^a\right|<\frac{\varepsilon}{5}.$$

The last estimate holds by (14):

$$\left|\bar{\mathbb{E}}g_{rk}^a-\lim_{j\to\infty}\bar{\mathbb{E}}g_{jk}^a\right|\leq C\eta(r)<\frac{\varepsilon}{5}.$$

These estimates combined, yield $\left|n^{-1}\sum_{i=0}^{n-1}\mathbb{E}g_{ik}^a-\lim_{j\to\infty}\bar{\mathbb{E}}g_{jk}^a\right|<\varepsilon$ for all $n\geq n_0$. Since $\lim_{j\to\infty}\bar{\mathbb{E}}g_{jk}^a$ exists, then also $\lim_{i\to\infty}n^{-1}\sum_{i=0}^{n-1}\mathbb{E}g_{ik}^a$ exists and is equal to it. □

Theorem 3.3 yields the next result as a corollary.

Theorem 3.6 *The series*

$$\sigma^2=\sum_{k=0}^{\infty}V_k,$$

where

$$V_k=(2-\delta_{k0})\lim_{r\to\infty}\bar{\mathbb{E}}[\mu(f_r f_{r+k})-\mu(f_r)\mu(f_{r+k})],$$

is absolutely convergent, and

$$\lim_{n\to\infty}\mathbb{E}\sigma_n^2=\sigma^2.$$

Proof Recall that $\mathbb{E}v_{ik}=(2-\delta_{k0})\mathbb{E}[\mu(f_i f_{i+k})-\mu(f_i)\mu(f_{i+k})]$. By Lemma 3.5 the limits $\lim_{n\to\infty}n^{-1}\sum_{i=0}^{n-1}\mathbb{E}\mu(f_i f_{i+k})=\lim_{j\to\infty}\bar{\mathbb{E}}\mu(f_j f_{j+k})$ and $\lim_{n\to\infty}n^{-1}\sum_{i=0}^{n-1}\mathbb{E}\mu(f_i)\mu(f_{i+k})=\lim_{j\to\infty}\bar{\mathbb{E}}\mu(f_j)\mu(f_{j+k})$ exist. Therefore

$$V_k=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}v_{ik}=(2-\delta_{k0})\lim_{r\to\infty}\bar{\mathbb{E}}[\mu(f_r f_{r+k})-\mu(f_r)\mu(f_{r+k})].$$

Now the rest of the claim follows from Theorem 3.3. □

Standing Assumption (*SA4*) for the rest of the paper we assume that $\mathbb{P}$ is asymptotically mean stationary, and there exist $C_0 > 0$ and $\zeta > 0$ such that

$$\sup_{r,k,a} \left| \frac{1}{n} \sum_{i=0}^{n-1} \int g_{rk}^a \circ \tau^i \, \mathrm{d}\mathbb{P} - \int g_{rk}^a \, \mathrm{d}\bar{\mathbb{P}} \right| \leq C_0 n^{-\zeta} \tag{15}$$

for all $n \geq 1$. Here the sup is taken over all $r, k \geq 0$ and $a \in \{1, 2\}$. ■

Lemma 3.7 *For all integers* $0 < n_1 < n_2$,

$$\left| (n_2 - n_1)^{-1} \sum_{i=n_1}^{n_2-1} \mathbb{E} g_{ik}^a - \lim_{r \to \infty} \bar{\mathbb{E}} g_{rk}^a \right| \leq C \left(\eta(n_1) + (n_2 - n_1)^{-\zeta} \right),$$

where C *is uniform.*

Proof By Lemma 3.4 we have

$$\begin{aligned} &\left| (n_2 - n_1)^{-1} \sum_{i=n_1}^{n_2-1} \mathbb{E} g_{ik}^a - (n_2 - n_1)^{-1} \sum_{i=n_1}^{n_2-1} \mathbb{E} \left[g_{n_1 k}^a \circ \tau^{i-n_1} \right] \right| \\ &\leq (n_2 - n_1)^{-1} \sum_{i=n_1}^{n_2-1} C\eta(n_1) = C\eta(n_1). \end{aligned} \tag{16}$$

By (15), it follows that

$$\begin{aligned} \left| (n_2 - n_1)^{-1} \sum_{i=n_1}^{n_2-1} \mathbb{E} \left[g_{n_1 k}^a \circ \tau^{i-n_1} \right] - \bar{\mathbb{E}} g_{n_1 k}^a \right| &= \left| (n_2 - n_1)^{-1} \sum_{i=0}^{n_2-n_1-1} \mathbb{E} \left[g_{n_1 k}^a \circ \tau^{i} \right] - \bar{\mathbb{E}} g_{n_1 k}^a \right| \\ &\leq C_0 (n_2 - n_1)^{-\zeta} \end{aligned} \tag{17}$$

Finally (14) gives

$$\left| \bar{\mathbb{E}} g_{n_1 k}^a - \lim_{r \to \infty} \bar{\mathbb{E}} g_{rk}^a \right| \leq C\eta(n_1). \tag{18}$$

Now the claim follows from (16), (17) and (18). □

Next we use Lemma 3.7 to provide an upper bound on $\left| n^{-1} \sum_{i=0}^{n-1} \mathbb{E} g_{ik}^a - \lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a \right|$. Note that just making the substitutions $n_1 = 0$ and $n_2 = n$ in Lemma 3.7 does not yield a good result. Instead we divide the sum $\sum_{i=0}^{n-1} \mathbb{E} g_{ik}^a$ into an increasing number of partial sums and then apply Lemma 3.7 separately to those parts.

Before proceeding to the next lemma, we define a function $h_\zeta : \mathbb{N} \to \mathbb{R}$ which depends on the parameter ζ in the following way

$$h_\zeta(n) = \begin{cases} n^{-1}, & \zeta > 1, \\ n^{-1} \log n, & \zeta = 1, \\ n^{-\zeta}, & 0 < \zeta < 1. \end{cases} \tag{19}$$

Lemma 3.8 *Suppose* $\eta(n) = Cn^{-\psi}$, $\psi > 1$. *Then a uniform bound*

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E} g_{ik}^a - \lim_{r \to \infty} \bar{\mathbb{E}} g_{rk}^a \right| \leq C h_\zeta(n)$$

holds.

Proof Denote $n^* = \lfloor \log_2 n \rfloor$. We split the sum $\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E} g_{ik}^a$ as

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E} g_{ik}^a = \frac{1}{n}\mathbb{E} g_{0k}^a + \frac{1}{n}\sum_{j=0}^{n^*-1}\sum_{i=2^j}^{2^{j+1}-1} \mathbb{E} g_{ik}^a + \frac{1}{n}\sum_{i=2^{n^*}}^{n-1} \mathbb{E} g_{ik}^a.$$

Obviously

$$\left|\frac{1}{n}\mathbb{E} g_{0k}^a - \frac{1}{n}\lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right| \le Cn^{-1}. \tag{20}$$

Lemma 3.7 yields

$$\begin{aligned}&\left|\sum_{i=2^j}^{2^{j+1}-1} \mathbb{E} g_{ik}^a - (2^{j+1} - 2^j)\lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right| \\ &\quad\le 2^j C((2^j)^{-\psi} + (2^{j+1} - 2^j)^{-\zeta}) \le C(2^{j(1-\psi)} + 2^{j(1-\zeta)}).\end{aligned}$$

Therefore

$$\begin{aligned}&\left|\frac{1}{n}\sum_{j=0}^{n^*-1}\sum_{i=2^j}^{2^{j+1}-1} \mathbb{E} g_{ik}^a - \frac{1}{n}(2^{n^*} - 1)\lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right| \\ &= \frac{1}{n}\left|\sum_{j=0}^{n^*-1}\left(\sum_{i=2^j}^{2^{j+1}-1} \mathbb{E} g_{ik}^a - (2^{j+1} - 2^j)\lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right)\right| \\ &\le Cn^{-1}\sum_{j=0}^{n^*-1}(2^{j(1-\psi)} + 2^{j(1-\zeta)}) \le C(n^{-1} + h_\zeta(n)) \le Ch_\zeta(n).\end{aligned} \tag{21}$$

Lemma 3.7 also gives

$$\begin{aligned}&\left|\frac{1}{n}\sum_{i=2^{n^*}}^{n-1} \mathbb{E} g_{ik}^a - \frac{1}{n}(n - 2^{n^*})\lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right| \\ &= n^{-1}(n - 2^{n^*})\left|(n - 2^{n^*})^{-1}\sum_{i=2^{n^*}}^{n-1} \mathbb{E} g_{ik}^a - \lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right| \\ &\le n^{-1}(n - 2^{n^*})C((2^{n^*})^{-\psi} + (n - 2^{n^*})^{-\zeta}) \le C(n^{-1} + h_\zeta(n)) \le Ch_\zeta(n).\end{aligned} \tag{22}$$

In the last line we used the fact that $n/2 \le 2^{n^*} \le n$, implying $n - 2^{n^*} \le n/2$. Collecting the estimates (20), (21) and (22), we deduce $\left|\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E} g_{ik}^a - \lim_{r\to\infty} \bar{\mathbb{E}} g_{rk}^a\right| \le Ch_\zeta(n)$. □

We are finally ready to state and prove the main result of this section:

Theorem 3.9 *Assume* (SA1) *and* (SA3) *with* $\eta(n) = Cn^{-\psi}$, $\psi > 1$. *Assume* (SA4) *with* $\zeta > 0$. *Then*

$$\left|\mathbb{E}\sigma_n^2 - \sigma^2\right| \le C\begin{cases} n^{\frac{1}{\psi}-1}, & \zeta > 1, \\ (n\log^{-1} n)^{\frac{1}{\psi}-1}, & \zeta = 1, \\ n^{\frac{\zeta}{\psi}-\zeta}, & 0 < \zeta < 1. \end{cases}$$

Here σ^2 *is the quantity appearing in Theorem 3.6.*

Proof Let $k \ge 0$. The previous lemma applied to case $a = 1$ yields

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[\mu(f_i f_{i+k})] - \lim_{r\to\infty} \bar{\mathbb{E}}[\mu(f_r f_{r+k})] \right| \le C h_\zeta(n). \tag{23}$$

Similarly in the case $a = 2$

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[\mu(f_i)(f_{i+k})] - \lim_{r\to\infty} \bar{\mathbb{E}}[\mu(f_r)\mu(f_{r+k})] \right| \le C h_\zeta(n). \tag{24}$$

Equations (23), (24) and Theorem 3.6 imply that

$$\begin{aligned}
&\left| V_k - \frac{1}{n} \sum_{i=0}^{n-1} (2 - \zeta_{k0}) \mathbb{E}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})] \right| \\
&\le C \left| \lim_{r\to\infty} \bar{\mathbb{E}}[\mu(f_r f_{r+k}) - \mu(f_r)\mu(f_{r+k})] - \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})] \right| \\
&\le C h_\zeta(n).
\end{aligned}$$

We apply Theorem 3.3, which yields

$$\left| \mathbb{E}\sigma_n^2 - \sigma^2 \right| \le C \left(\left| \sum_{k=0}^{K} h_\zeta(n) \right| + \sum_{k=K+1}^{\infty} k^{-\psi} + K n^{-1} \right) \le C K (h_\zeta(n) + K^{-\psi}), \tag{25}$$

for all $0 < K \le n$. The estimate on the right side of (25) is minimized, when $h_\zeta(n) = K^{-\psi}$. Therefore choosing

$$K \asymp \begin{cases} n^{\frac{1}{\psi}}, & \zeta > 1, \\ (n \log^{-1} n)^{\frac{1}{\psi}}, & \zeta = 1, \\ n^{\frac{\zeta}{\psi}}, & 0 < \zeta < 1, \end{cases}$$

in (25) completes the proof. □

4 Conclusions

4.1 Main Result and Consequences

Theorems 2.14 and 3.9 immediately yield the main result of the paper, given next. The bounds shown are elementary combinations of these theorems, so we leave the details to the reader. Let us remind the reader of the Standing Assumptions (SA1)–(SA4) in Sects. 2 and 3 . At the end of the section we also comment on the case of vector-valued observables.

Theorem 4.1 *Assume* (SA1 and 3) *with* $\eta(n) = Cn^{-\psi}$, $\psi > 1$; (SA2) *with* $\alpha(n) = Cn^{-\gamma}$, $\gamma > 0$; *and* (SA4) *with* $\zeta > 0$. *Fix an arbitrarily small* $\delta > 0$. *Then there exists* $\Omega_* \subset \Omega$, $\mathbb{P}(\Omega_*) = 1$, *such that all of the following holds*: *the non-random number*[3]

$$\sigma^2 = \sum_{k=0}^{\infty} (2 - \delta_{k0}) \lim_{i\to\infty} \bar{\mathbb{E}}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})]$$

[3] Here $\delta_{k0} = 1$ if $k = 0$, and $\delta_{k0} = 0$ if $k \neq 0$.

is well defined, nonnegative, the series is absolutely convergent, and

$$\lim_{n\to\infty} \sigma_n^2(\omega) = \sigma^2$$

for every $\omega \in \Omega_*$. *Moreover, the absolute difference*

$$\Delta_n(\omega) = \left|\sigma_n^2(\omega) - \sigma^2\right|$$

has the following upper bounds, for any $\omega \in \Omega_*$:

$$\Delta_n = \begin{cases} O\left(n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n\right), & \zeta \ge 1,\ \min\{\psi-1,\gamma\} > 1,\\ O\left(n^{-\frac{1}{2}+\delta}\right), & \zeta \ge 1,\ \min\{\psi-1,\gamma\} = 1,\\ O\left(n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n\right), & \zeta \ge 1,\ 0 < \min\{\psi-1,\gamma\} < 1,\\ O\left(n^{\frac{\zeta}{\psi}-\zeta} + n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n\right), & 0<\zeta<1,\ \min\{\psi-1,\gamma\} > 1,\\ O\left(n^{\frac{\zeta}{\psi}-\zeta} + n^{-\frac{1}{2}+\delta}\right), & 0<\zeta<1,\ \min\{\psi-1,\gamma\} = 1,\\ O\left(n^{\frac{\zeta}{\psi}-\zeta} + n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n\right), & 0<\zeta<1,\ 0 < \min\{\psi-1,\gamma\} < 1. \end{cases}$$

Let us reiterate that Theorem 4.1 facilitates proving quenched central limit theorems with convergence rates for the fiberwise centered $\bar{W}_n$. Recalling the discussion from the beginning of the paper, we namely have the following trivial lemma (thus presented without proof):

Lemma 4.2 *Suppose* $d(\cdot\,,\,\cdot)$ *is a distance of probability distributions with the following property: given* $b > 0$, *there exist an open neighborhood* $U \subset \mathbb{R}_+$ *of* b *and a constant* $C > 0$, *such that*

$$d(aZ, bZ) \le C|a-b| \tag{26}$$

for all $a \in U$. *Here* $Z \sim \mathcal{N}(0,1)$. *If* $\sigma^2 > 0$, *then for every* $\omega \in \Omega_*$,

$$d(\bar{W}_n, \sigma Z) \le d(\bar{W}_n, \sigma_n Z) + O(\Delta_n).$$

In other words, once a bound on the first term on the right side has been established (e.g., using methods cited earlier), one can use Theorem 4.1 to bound the second term almost surely. Typical metrics satisfying (26) are the 1-Lipschitz (Wasserstein) and Kolmogorov distances.

The results presented above allow to formulate some sufficient conditions for $\sigma^2 > 0$. For simplicity, we proceed in the ideal parameter regime

$$\min\{\psi-1, \gamma, \zeta\} > 1. \tag{27}$$

Generalizations of the next result involving any of the other parameter regimes of Theorem 4.1 are straightforward, and left to the reader.

Corollary 4.3 *Let* (27) *hold with all the other assumptions of Theorem 4.1. Suppose that either* (i) *there exists* $\omega \in \Omega_*$ *such that*

$$\sup_{n\ge 2} \frac{\operatorname{Var}_\mu(S_n)}{n^{\frac{1}{2}} \log^{\frac{3}{2}+\delta} n} = \infty$$

or (ii) *and*

$$\sup_{n\ge 1} \frac{\mathbb{E}\operatorname{Var}_\mu(S_n)}{n^{\frac{1}{\psi}}} = \infty.$$

Then $\sigma^2 > 0$.

Proof Suppose $\sigma^2 = 0$. We will derive a contradiction in each case.
(i) Let $\omega \in \Omega_*$ be arbitrary. By Theorem 4.1, there exists $C > 0$ such that $n^{-1}\operatorname{Var}_\mu(S_n) = \sigma_n^2 \le Cn^{-\frac{1}{2}}\log^{\frac{3}{2}+\delta} n$ for all $n \ge 1$. This violates the assumption of part (i), so $\sigma^2 > 0$.
(ii) By Theorem 3.9, there exists $C > 0$ such that $n^{-1}\mathbb{E}\operatorname{Var}_\mu(S_n) = \mathbb{E}\sigma_n^2 \le Cn^{\frac{1}{\psi}-1}$ for all $n \ge 1$. This violates the assumption of part (ii), so $\sigma^2 > 0$. ◻

We will return to the question of whether $\sigma^2 = 0$ or $\sigma^2 > 0$ in Lemma B.1.

4.2 Vector-Valued Observables

Let us conclude by explaining, as promised, how the results extend with ease to the case of a vector-valued observable $f : X \to \mathbb{R}^d$. This time σ_n^2 is a $d \times d$ covariance matrix and, if the limit exists, so is $\sigma^2 = \lim_{n\to\infty}\sigma_n^2$. Define the functions $\ell_n : \mathbb{R}^d \to \mathbb{R}$ by

$$\ell_n(v) = v^T\sigma_n^2 v,$$

and denote the standard basis vectors of $\mathbb{R}^d$ by e_α, $\alpha = 1, \ldots, d$. Observe that $\ell_n(v)$ is the μ-variance of W_n with the scalar-valued observable $v^T f$ in place of f.

Lemma 4.4 *Suppose there exists $\kappa > 0$ such that, almost surely, the limit $\ell(e_\alpha + e_\beta) = \lim_{n\to\infty}\ell_n(e_\alpha + e_\beta)$ exists and*

$$\ell(e_\alpha + e_\beta) - \ell_n(e_\alpha + e_\beta) = O(n^{-\kappa})$$

as $n \to \infty$ for all $\alpha, \beta = 1, \ldots, d$. Then, almost surely, $\sigma^2 = \lim_{n\to\infty}\sigma_n^2$ exists and

$$\left|\sigma^2 - \sigma_n^2\right| = O(n^{-\kappa})$$

for all matrix norms.

Proof Note that the matrix elements of σ_n^2 are given by

$$\left(\sigma_n^2\right)_{\alpha\beta} = \tfrac{1}{2}(\ell_n(e_\alpha + e_\beta) - \ell_n(e_\alpha) - \ell_n(e_\beta)).$$

Dropping the subindex n yields the limit matrix elements $\sigma_{\alpha\beta}^2$. Since α and β can take only finitely many values, simultaneous almost sure convergence for the matrix elements with the claimed rate follows. ◻

According to the lemma, the rate of convergence of the covariance matrix σ_n^2 to σ^2 can be established by applying the earlier results to the finite family of scalar-valued observables $(e_\alpha + e_\beta)^T f$. Further, one may apply Corollary 4.3 (or Lemma B.1) to the observables $v^T f$ for all unit vectors v to obtain conditions for σ^2 being positive definite. Assuming now it is, for certain metrics (e.g. 1-Lipschitz) one has

$$d(\sigma_n Z, \sigma Z) \le C\left|\sigma^2 - \sigma_n^2\right|$$

where $Z \sim \mathcal{N}(0, I_{d\times d})$ and $C = C(\sigma)$, which again yields an estimate of the type

$$d(\bar{W}_n, \sigma Z) \le d(\bar{W}_n, \sigma_n Z) + C\left|\sigma^2 - \sigma_n^2\right|.$$

We refer the reader to Hella [13] for details, including the hard part of establishing an almost sure, asymptotically decaying bound on $d(\bar{W}_n, \sigma_n Z)$ in the vector-valued case.

Remark 4.5 As an application, Hella [13] establishes the convergence rate $n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n$ for random compositions of uniformly expanding circle maps in the regime (27). Furthermore, Leppänen and Stenlund [16] establish the same result for random compositions of non-uniformly expanding Pomeau–Manneville maps.

Acknowledgements Open access funding provided by University of Helsinki including Helsinki University Central Hospital.

Appendix A. Random Dynamical Systems

In this section we interpret the limit variance of Theorems 3.6 and 3.9 from the point of view of RDSs. Like elsewhere in the paper, we will assume the system possesses the good, uniform, fiberwise properties of the Standing Assumptions.

Recall that τ preserves the probability measure $\bar{\mathbb{P}}$ in (12), i.e., $\tau^{-1}F \in \mathcal{F}$ and $\bar{\mathbb{P}}(\tau^{-1}F) = \bar{\mathbb{P}}(F)$ for all $F \in \mathcal{F}$. One says that $\varphi(\,\cdot\,,\,\cdot\,,\,\cdot\,)$ in (1) is a measurable RDS on the measurable space $(X, \mathcal{B})$ over the measure-preserving transformation $(\Omega, \mathcal{F}, \bar{\mathbb{P}}, \tau)$. The map

$$\Phi : \Omega \times X \to \Omega \times X : \Phi(\omega, x) = (\tau\omega, \varphi(1, \omega)x) = (\tau\omega, T_{\omega_1}(x))$$

is called the skew product of the measure-preserving transformation $(\Omega, \mathcal{F}, \bar{\mathbb{P}}, \tau)$ and the cocycle $\varphi(n, \omega)$ on X. It is a measurable self-map on $(\Omega \times X, \mathcal{F} \otimes \mathcal{B})$. In general, RDSs and skew products have one-to-one correspondence; in particular, the measurability of one implies the measurability of the other.

We are interested in the skew product, because of the identity

$$\Phi^n(\omega, x) = (\tau^n\omega, \varphi(n, \omega)x) = (\tau^n\omega, T_{\omega_n} \circ \cdots \circ T_{\omega_1}(x)).$$

Thus, our task is to study the statistics of the projection of $\Phi^n(\omega, x)$ to X. It now becomes interesting to study the invariant measures of Φ. However, the class of all invariant measures of Φ is unnatural, for we must incorporate the fact that τ preserves the measure $\bar{\mathbb{P}}$. For this reason, it is said that a probability measure $\mathbf{P}$ on $\mathcal{F} \otimes \mathcal{B}$ is an invariant measure for the RDS φ if it is invariant for Φ *and* the marginal of $\mathbf{P}$ on Ω coincides with $\bar{\mathbb{P}}$. In other words,

$$\Phi_*\mathbf{P} = \mathbf{P} \quad \text{and} \quad (\Pi_1)_*\mathbf{P} = \bar{\mathbb{P}},$$

where $\Pi_1 : \Omega \times X \to X : (\omega, x) \mapsto \omega$.

We will also need to consider the cocycle

$$\varphi^{(2)}(n, \omega)(x, y) = (\varphi(n, \omega)x, \varphi(n, \omega)y)$$

on the product space $X \times X$. The corresponding skew product is

$$\Phi^{(2)}(\omega, x, y) = (\tau\omega, \varphi^{(2)}(1, \omega)(x, y)).$$

The invariant measures of the RDS $\varphi^{(2)}$ are defined analogously to above. Without danger of confusion, we define the projections $\Pi_1(\omega, x, y) = \omega$, $\Pi_2(\omega, x, y) = x$ and $\Pi_3(\omega, x, y) = y$ on $\Omega \times X \times X$. We also write $(\Pi_1 \times \Pi_2)(\omega, x, y) = (\omega, x)$ and $(\Pi_1 \times \Pi_3)(\omega, x, y) = (\omega, y)$.

Of particular interest will be the sequence of functions $Z_n : \Omega \times X \times X$ defined by

$$Z_n(\omega, x, y) = S_n(\omega, x) - S_n(\omega, y).$$

For then

$$\int Z_n^2(\omega)\,\mathrm{d}(\mu \otimes \mu) = 2\,\mathrm{Var}_\mu(S_n(\omega)) = 2\sigma_n^2(\omega) \cdot n. \tag{28}$$

Notice already that writing

$$F(\omega, x, y) = f(x) - f(y)$$

yields the identity

$$Z_n = \sum_{i=0}^{n-1} F \circ (\Phi^{(2)})^i. \tag{29}$$

Standing Assumption (*SA5*) assume there exists an invariant measure $\mathbf{P}^{(2)}$ for the RDS $\varphi^{(2)}$ that is symmetric in the sense that

$$\int h(\omega, x, y)\,\mathrm{d}\mathbf{P}^{(2)}(\omega, x, y) = \int h(\omega, y, x)\,\mathrm{d}\mathbf{P}^{(2)}(\omega, x, y) \tag{30}$$

for all bounded measurable $h : \Omega \times X \times X \to \mathbb{R}$. The common marginal

$$\mathbf{P} = (\Pi_1 \times \Pi_2)_* \mathbf{P}^{(2)} = (\Pi_1 \times \Pi_3)_* \mathbf{P}^{(2)} \tag{31}$$

is then trivially an invariant measure for the RDS φ. Moreover, assume

$$\lim_{i \to \infty} \bar{\mathbb{E}}[\mu(f_i)] = \int f(x)\,\mathrm{d}\mathbf{P}(\omega, x), \tag{32}$$

$$\lim_{i \to \infty} \bar{\mathbb{E}}[\mu(f_i f_{i+k})] = \int f(x)\, f(\varphi(k, \omega, x))\,\mathrm{d}\mathbf{P}(\omega, x) \tag{33}$$

and

$$\lim_{i \to \infty} \bar{\mathbb{E}}[\mu(f_i)\mu(f_{i+k})] = \int f(x)\, f(\varphi(k, \omega, y))\,\mathrm{d}\mathbf{P}^{(2)}(\omega, x, y) \tag{34}$$

are satisfied. ■

While Standing Assumption (SA5) may, from the point of view of the initial setup of our problem, seem mysterious at a first glance, it is quite natural. We will later provide an example of a more concrete condition which implies (SA5), and stick to the abstract setting for now.

The following lemma lists useful properties of F in view of (SA5).

Lemma A.1 *The function F satisfies*

$$\int F\,\mathrm{d}\mathbf{P}^{(2)} = 0$$

and

$$\int F \cdot F \circ (\Phi^{(2)})^i\,\mathrm{d}\mathbf{P}^{(2)} = 2 \int f(x) f(\varphi(i, \omega, x)) - f(x) f(\varphi(i, \omega, y))\,\mathrm{d}\mathbf{P}^{(2)}(\omega, x, y). \tag{35}$$

The latter has the upper bound

$$\left| \int F \cdot F \circ (\Phi^{(2)})^i \, \mathrm{d}\mathbf{P}^{(2)} \right| \leq 2\eta(i). \tag{36}$$

Proof That F is centered is due to the symmetry property (30) of $\mathbf{P}^{(2)}$ in (SA5). Since

$$\left(F \cdot F \circ (\Phi^{(2)})^i\right)(\omega, x, y) = \{f(x) - f(y)\}\{f(\varphi(i, \omega, x)) - f(\varphi(i, \omega, y))\},$$

the same symmetry property also yields (35). The upper bound in (36) then follows from (33) and (34) in (SA5) together with (SA1). □

Recall that in Theorems 3.6, 3.9 and 4.1 we have

$$\sigma^2 = \lim_{n\to\infty} \mathbb{E}\sigma_n^2 = \sum_{k=0}^{\infty}(2 - \delta_{k0}) \lim_{i\to\infty} \bar{\mathbb{E}}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})].$$

The next lemma connects this expression to the RDS notions when also (SA5) is assumed.

Lemma A.2 *The limit variance σ^2 in Theorems 3.6, 3.9 and 4.1 satisfies*

$$\sigma^2 = \sum_{k=0}^{\infty}(2 - \delta_{k0}) \int f(x) f(\varphi(k, \omega, x)) - f(x) f(\varphi(k, \omega, y)) \, \mathrm{d}\mathbf{P}^{(2)}(\omega, x, y) \tag{37}$$

$$= \frac{1}{2} \sum_{k=0}^{\infty}(2 - \delta_{k0}) \int F \cdot F \circ (\Phi^{(2)})^k \, \mathrm{d}\mathbf{P}^{(2)} \tag{38}$$

$$= \frac{1}{2} \lim_{n\to\infty} \frac{1}{n} \int Z_n^2 \, \mathrm{d}\mathbf{P}^{(2)} = \frac{1}{2} \lim_{n\to\infty} \mathrm{Var}_{\mathbf{P}^{(2)}}\left(\frac{Z_n}{\sqrt{n}}\right). \tag{39}$$

Proof The first line is just the expression of σ^2 rewritten using (33) and (34). The second line then follows by (35). The last line holds by (29) together with (36) and (SA1). □

Remark A.3 Note that the expression of σ^2 in (38) is exactly **one half** of the Green–Kubo formula in terms of the skew-product $\Phi^{(2)}$, its invariant measure $\mathbf{P}^{(2)}$, and the observable F. This trick of "doubling the dimension" is not new. To our knowledge, however, (38) is a new observation at this level of generality. It answers a question raised in [2, Sect. 7] by Aimino, Nicol and Vaienti (who studied the special case where $\mathbb{P}$, $\mathbf{P}$ and $\mathbf{P}^{(2)}$ are product measures, allowing for a non-random centering of S_n): The key that makes (38) an algebraic fact is the **symmetry** property (30) of the measure $\mathbf{P}^{(2)}$.

It deserves a separate remark that even though σ^2 does not in general (see Remark C.2) admit a classical Green–Kubo formula in terms of Φ, $\mathbf{P}$, and f, "doubling the dimension" still yields (38).

Appendix B. Positivity of σ^2

In this section we return to the question of positivity of the limit variance σ^2. We shall assume (SA1) and (SA3)–(SA5), the strong-mixing assumption (SA2) being unnecessary here. Again we assume nice parameters—e.g. $\psi > 2$—for simplicity of the statements.

The foregoing discussion allows us to give some characterizations of the cases $\sigma^2 = 0$ and $\sigma^2 > 0$ on various levels of abstraction:

Lemma B.1 *Suppose $\eta(n) = Cn^{-\psi}$, $\psi > 2$.*

(i) $\sigma^2 = 0$ is equivalent to each of the following conditions:

(a) $\sup_{n\geq 0} \int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)} < \infty$.
(b) *There exists $G \in L^2(\mathbf{P}^{(2)})$ such that $F = G - G \circ \Phi^{(2)}$.*

(ii) *$\sigma^2 > 0$ is equivalent to each of the following conditions:*

(a) $\sup_{n\geq 0} \int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)} = \infty$.
(b) *There exist $c > 0$ and $N > 0$ such that $\int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)} \geq cn$ for all $n \geq N$.*

(iii) *If $\zeta > 1$, then $\sigma^2 > 0$ is equivalent to each of the following conditions:*

(a) $\sup_{n\geq 1} n^{-\frac{1}{\psi}} \int Z_n^2 \,\mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) = \infty$.
(b) $\sup_{n\geq 1} n^{-\frac{1}{\psi}} \,\mathbb{E}\,\mathrm{Var}_\mu(S_n) = \infty$.
(c) *There exist $c > 0$ and $N > 0$ such that $\int Z_n^2 \,\mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) \geq cn$ for all $n \geq N$.*
(d) *There exist $c > 0$ and $N > 0$ such that $\mathbb{E}\,\mathrm{Var}_\mu(S_n) \geq cn$ for all $n \geq N$.*

(iv) *If $\mathbb{P}$ is stationary, then $\sigma^2 = 0$ is equivalent to each of the following conditions:*

(a) $\sup_{n\geq 1} \int Z_n^2 \,\mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) < \infty$.
(b) $\sup_{n\geq 1} \mathbb{E}\,\mathrm{Var}_\mu(S_n) < \infty$.

(v) *If $\mathbb{P}$ is stationary, then $\sigma^2 > 0$ is equivalent to each of the following conditions:*

(a) $\sup_{n\geq 1} \int Z_n^2 \,\mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) = \infty$.
(b) $\sup_{n\geq 1} \mathbb{E}\,\mathrm{Var}_\mu(S_n) = \infty$.
(c) *There exist $c > 0$ and $N > 0$ such that $\int Z_n^2 \,\mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) \geq cn$ for all $n \geq N$.*
(d) *There exist $c > 0$ and $N > 0$ such that $\mathbb{E}\,\mathrm{Var}_\mu(S_n) \geq cn$ for all $n \geq N$.*

From the point of view of applications, parts (iii)(b and d), (iv)(b) and (v)(b and d) may be the most relevant ones as they involve the measures $\mathbb{P}$ and μ, and the process $(S_n)_{n\geq 1}$, which are immediately apparent from the definition of the system. Note that (iii)(b) is the same condition as in Corollary 4.3(ii).

Proof of Lemma B.1 By (36) we can appeal to a well-known result due to Leonov [15], which guarantees that the limit $b = \lim_{n\to\infty} \int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)}$ exists in $[0, \infty]$, and $b < \infty$ if and only if $\sup_{n\geq 0} \int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)} < \infty$. Moreover, the last condition is equivalent to the existence of $G \in L^2(\mathbf{P}^{(2)})$ such that $F = G - G \circ \Phi^{(2)}$. On the other hand, standard computations and the formula for σ^2 in (38) yield

$$\begin{aligned}\int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)} &= 2\sigma^2 n - 2n \sum_{k=n}^{\infty} \int F \cdot F \circ (\Phi^{(2)})^k \,\mathrm{d}\mathbf{P}^{(2)} - 2 \sum_{k=1}^{n-1} k \int F \cdot F \circ (\Phi^{(2)})^k \,\mathrm{d}\mathbf{P}^{(2)} \\ &= 2\sigma^2 n + O\left(n \sum_{k=n}^{\infty} k^{-\psi} + \sum_{k=1}^{n-1} k^{1-\psi} \right) = 2\sigma^2 n + O(1).\end{aligned}$$

Here $\psi > 2$ was used. Thus, $\sigma^2 > 0$ is equivalent to linear growth of $\int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)}$ to infinity, while $\sigma^2 = 0$ is equivalent to $\sup_{n\geq 0} \int Z_n^2 \,\mathrm{d}\mathbf{P}^{(2)} < \infty$. Parts (i) and (ii) are proved.

As for part (iii), (28) and Theorem 3.9 with $\zeta > 1$ yield

$$\int Z_n^2 \,\mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) = 2\,\mathbb{E}\,\mathrm{Var}_\mu(S_n) = 2\left(\sigma^2 + O\left(n^{\frac{1}{\psi}-1}\right)\right) n = 2\sigma^2 n + O\left(n^{\frac{1}{\psi}}\right).$$

If $\sigma^2 = 0$, the right side of the estimate is $O\left(n^{\frac{1}{\psi}}\right)$, and each of the conditions (a)–(d) fails. If $\sigma^2 > 0$, the right side grows asymptotically linearly in n, and (a)–(d) are all satisfied.

Finally, parts (iv) and (v) follow from (i) and (ii), respectively, because in the stationary case it holds that $\int Z_n^2 \, \mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) = \int Z_n^2 \, \mathrm{d}\mathbf{P}^{(2)} + O(1)$; see Lemma B.2 below. □

We close the section with the lemma below, which was needed in the last part of the preceding proof.

Lemma B.2 *Suppose $\mathbb{P}$ is stationary and $\eta(n) = Cn^{-\psi}$, $\psi > 2$. Then*

$$\sup_{n\ge 1}\left|\int Z_n^2 \, \mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) - \int Z_n^2 \, \mathrm{d}\mathbf{P}^{(2)}\right| < \infty.$$

Proof We have

$$\int Z_n^2 \, \mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) = \sum_{i=0}^{n-1}\sum_{k=0}^{n-i-1}(2-\delta_{k0})\int F\circ(\Phi^{(2)})^i \, F\circ(\Phi^{(2)})^{i+k} \, \mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) \tag{40}$$

and

$$\begin{aligned}\int Z_n^2 \, \mathrm{d}\mathbf{P}^{(2)} &= \sum_{i=0}^{n-1}\sum_{k=0}^{n-i-1}(2-\delta_{k0})\int F\circ(\Phi^{(2)})^i \, F\circ(\Phi^{(2)})^{i+k} \, \mathrm{d}\mathbf{P}^{(2)} \\ &= \sum_{i=0}^{n-1}\sum_{k=0}^{n-i-1}(2-\delta_{k0})\int F\cdot F\circ(\Phi^{(2)})^k \, \mathrm{d}\mathbf{P}^{(2)}.\end{aligned} \tag{41}$$

Denote

$$a_{ik} = \int F\circ(\Phi^{(2)})^i \, F\circ(\Phi^{(2)})^{i+k} \, \mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) = 2\mathbb{E}[\mu(f_i f_{i+k})] - 2\mathbb{E}[\mu(f_i)\mu(f_{i+k})]$$

and

$$b_k = \int F\cdot F\circ(\Phi^{(2)})^k \, \mathrm{d}\mathbf{P}^{(2)}.$$

Note that $|a_{ik}| \le 2\eta(k)$ by (SA1) and $|b_k| \le 2\eta(k)$ by (36). Thus $|a_{ik} - b_k| \le Ck^{-\psi}$ for all i and k. By stationarity ($\mathbb{P} = \bar{\mathbb{P}}$) and (SA5),

$$\lim_{i\to\infty} a_{ik} = b_k.$$

Again by stationarity, (14) implies that $|a_{ik} - b_k| \le C\eta(i) = Ci^{-\psi}$. Thus

$$|a_{ik} - b_k| \le C\max\{i,k\}^{-\psi}. \tag{42}$$

Collecting (40), (41) and (42) we get the estimate

$$\begin{aligned}&\left|\int Z_n^2 \, \mathrm{d}(\mathbb{P} \otimes \mu \otimes \mu) - \int Z_n^2 \, \mathrm{d}\mathbf{P}^{(2)}\right| \\ &\le \sum_{i=0}^{n-1}\sum_{k=0}^{n-i-1}(2-\delta_{k0})|a_{ik} - b_k| \le C\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\max\{i,k\}^{-\psi} = C\sum_{k=0}^{\infty}(2k+1)k^{-\psi} < \infty\end{aligned}$$

for all n. The proof is complete. □

Appendix C. Effect of the Fiberwise Centering of W_n

In this section we discuss Remark 1.2 concerning the variance of W_n, as opposed to the fiberwise-centered $\bar{W}_n = W_n - \mu(W_n)$. Note that

$$\operatorname{Var}_{\mathbb{P}\otimes\mu} W_n = \mathbb{E}\mu\left(W_n^2\right) - (\mathbb{E}\mu(W_n))^2 \tag{43}$$

and

$$\operatorname{Var}_{\mathbb{P}} \mu(W_n) = \mathbb{E}[\mu(W_n)^2] - (\mathbb{E}\mu(W_n))^2. \tag{44}$$

The difference of (43) and (44) equals

$$\mathbb{E}\sigma_n^2 = \operatorname{Var}_{\mathbb{P}\otimes\mu} \bar{W}_n = \mathbb{E}\mu\left(W_n^2\right) - \mathbb{E}[\mu(W_n)^2].$$

Under the assumptions of our paper

$$\mathbb{E}\sigma_n^2 = \sigma^2 + o(1).$$

Therefore $\operatorname{Var}_{\mathbb{P}\otimes\mu} W_n$ and $\operatorname{Var}_{\mathbb{P}} \mu(W_n)$ either converge or diverge simultaneously. We now derive their asymptotic expressions in terms of series, restricting to the case where the law $\mathbb{P}$ of the selection process is **stationary**.

Lemma C.1 *Let $\mathbb{P}$ be stationary. Let* (SA1)–(SA3) *hold with $\eta(n) = Cn^{-\psi}$, $\psi > 1$, and $\alpha(n) = n^{-\gamma}$, $\gamma > 1$. Then*

$$\operatorname{Var}_{\mathbb{P}\otimes\mu} W_n = \sum_{k=0}^{\infty}(2-\delta_{k0}) \lim_{i\to\infty} \{\mathbb{E}[\mu(f_i f_{i+k})] - \mathbb{E}\mu(f_i)\mathbb{E}\mu(f_{i+k})\} + O\left(n^{\frac{1}{\min\{\gamma,\psi\}}-1}\right)$$

and

$$\operatorname{Var}_{\mathbb{P}} \mu(W_n) = \sum_{k=0}^{\infty}(2-\delta_{k0}) \lim_{i\to\infty} \{\mathbb{E}[\mu(f_i)\mu(f_{i+k})] - \mathbb{E}\mu(f_i)\mathbb{E}\mu(f_{i+k})\} + O\left(n^{\frac{1}{\min\{\gamma,\psi\}}-1}\right).$$

Here the limits exist and the series converge absolutely. If also (SA5) *holds, then*

$$\lim_{n\to\infty} \operatorname{Var}_{\mathbb{P}\otimes\mu} W_n = \sum_{k=0}^{\infty}(2-\delta_{k0}) \left\{\int f(x)\, f(\varphi(k,\omega,x))\, \mathrm{d}\mathbf{P}(\omega,x) - \left(\int f(x)\, \mathrm{d}\mathbf{P}(\omega,x)\right)^2\right\}$$

and

$$\lim_{n\to\infty} \operatorname{Var}_{\mathbb{P}} \mu(W_n) = \sum_{k=0}^{\infty}(2-\delta_{k0}) \left\{\int f(x)\, f(\varphi(k,\omega,y))\, \mathrm{d}\mathbf{P}^{(2)}(\omega,x,y) - \left(\int f(x)\, \mathrm{d}\mathbf{P}(\omega,x)\right)^2\right\}$$

using the RDS notations.

Remark C.2 Note that in the latter case

$$\lim_{n\to\infty} \operatorname{Var}_{\mathbb{P}\otimes\mu} W_n = \sum_{k=0}^{\infty}(2-\delta_{k0}) \operatorname{Cov}_{\mathbf{P}}(f, f\circ\Phi^k).$$

This is the classical Green–Kubo formula in terms of the skew-product Φ, its invariant measure $\mathbf{P}$, and the observable f. Let us stress that it is not the expression of σ^2, save for exactly the special case $\lim_{n\to\infty} \mathrm{Var}_{\mathbb{P}}\, \mu(W_n) = 0$. The latter special case is the very same in which Abdelkader and Aimino [1] establish a quenched central limit theorem with non-random centering, assuming i.i.d. randomness ($\mathbb{P} = \mathbb{P}_0^{\mathbb{N}}$) in particular; see also Remark A.3.

Proof of Lemma C.1 We prove the statements concerning $\mathrm{Var}_{\mathbb{P}}\, \mu(W_n)$ first. We have

$$\mathbb{E}\left[\mu(W_n)^2\right] - (\mathbb{E}\mu(W_n))^2 = \frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{n-i-1} a_{ik}$$

where

$$a_{ik} = (2-\delta_{k0})\{\mathbb{E}[\mu(f_i)\mu(f_{i+k})] - \mathbb{E}\mu(f_i)\mathbb{E}\mu(f_{i+k})\}.$$

We will apply Lemma 3.2 to show convergence as $n \to \infty$. To that end, we need control of a_{ik} in the limits $i \to \infty$ and $k \to \infty$. We begin with the first limit.

By (5) below (SA3), we have a uniform bound

$$|\mu(f\circ\varphi(i,\omega)) - \mu(f\circ\varphi(r,\tau^{i-r}\omega))| \le C\eta(r) \tag{45}$$

whenever $r \le i$. Since $\mathbb{P}$ is stationary, this yields

$$|\mathbb{E}\mu(f_i) - \mathbb{E}\mu(f_r)| \le C\eta(r).$$

Thus, $(\mathbb{E}\mu(f_i))_{i=0}^{\infty}$ is Cauchy, so its limit exists and

$$\left|\lim_{i\to\infty} \mathbb{E}\mu(f_i)\mathbb{E}\mu(f_{i+k}) - \mathbb{E}\mu(f_r)\mathbb{E}\mu(f_{r+k})\right| \le C\eta(r).$$

Since $\bar{\mathbb{P}} = \mathbb{P}$ by stationarity, (14) gives

$$\left|\lim_{i\to\infty} \mathbb{E}[\mu(f_i)\mu(f_{i+k})] - \mathbb{E}[\mu(f_r)\mu(f_{r+k})]\right| \le C\eta(r).$$

Thus, the limit

$$b_k = (2-\delta_{k0})\lim_{i\to\infty}\{\mathbb{E}[\mu(f_i)\mu(f_{i+k})] - \mathbb{E}\mu(f_i)\mathbb{E}\mu(f_{i+k})\}$$

exists and

$$a_{ik} = b_k + O(\eta(i))$$

as $i \to \infty$. Since η is summable,

$$r_k(n) = \frac{1}{n}\sum_{i=0}^{n-1} a_{ik} - b_k = \frac{1}{n}\sum_{i=0}^{n-1} O(\eta(i)) = O(n^{-1})$$

as $n \to \infty$. Both of the preceding bounds are uniform in k.

In order to bound a_{ik} as $k \to \infty$, first note that (45) allows to estimate

$$|\mu(f_{i+k}) - v| \le C\eta(k-r)$$

for $r \le k$, where the function $v(\omega) = \mu(f\circ\varphi(k-r,\tau^{i+r}\omega))$ is $\mathcal{F}_{i+r+1}^{\infty}$-measurable and bounded; see Sect. 2.1 for terminology. Thus

$$a_{ik} = (2-\delta_{k0})\{\mathbb{E}[\mu(f_i)v] - \mathbb{E}\mu(f_i)\mathbb{E}v\} + O(\eta(k-r)) = O(\alpha(r)) + O(\eta(k-r)),$$

the last estimate being true by strong mixing. Picking $r \asymp k/2$ yields

$$a_{ik} = O(k^{-\min\{\gamma,\psi\}})$$

uniformly in i. Since $\gamma > 1$ and $\psi > 1$, this bound is summable, so Lemma 3.2 can now be applied; recall (10). The bound in (11) becomes

$$\left| \mathbb{E}\left[\mu(W_n)^2\right] - (\mathbb{E}\mu(W_n))^2 - \sum_{k=0}^{\infty} b_k \right| \le C(K^{1-\min\{\gamma,\psi\}} + Kn^{-1}).$$

Now, choosing $K \asymp n^{1/\min\{\gamma,\psi\}}$ yields the upper bound $Cn^{1/\min\{\gamma,\psi\}-1}$ claimed.

The expressions of the limits b_k in term of the RDS notations is obtained with the help of (32)–(34), recalling again $\mathbb{P} = \bar{\mathbb{P}}$ due to stationarity.

Finally, the claims regarding $\operatorname{Var}_{\mathbb{P}\otimes\mu} W_n = \operatorname{Var}_{\mathbb{P}\otimes\mu} \bar{W}_n + \operatorname{Var}_{\mathbb{P}} \mu(W_n)$ follow since we already have control of both terms on the right side: in the stationary case at hand, Theorem 3.9 applies with any $\zeta > 1$, yielding $\operatorname{Var}_{\mathbb{P}\otimes\mu} \bar{W}_n = \sigma^2 + O\left(n^{\frac{1}{\psi}-1}\right)$. □

Appendix D. (SA5′): A Less Abstract Substitute for (SA5)

Standing Assumption (SA5) is abstract in that it involves the invariant measure $\mathbf{P}^{(2)}$ of the RDS $\varphi^{(2)}$, and a number of properties of the measure, which are not obvious from the setup of the system at the beginning of the paper. For that reason we give in this section, as an example, another assumption which (i) is more concrete in that it involves only the initial measure μ and the basic cocycle φ, and (ii) is stronger than (SA5).

Standing Assumption (*SA5′*) throughout this section we assume following: the measures $\varphi(n,\omega)_*\mu$ have uniformly square integrable densities with respect to μ, i.e., there exists $K > 0$ such that

$$\left\| \frac{\mathrm{d}\varphi(n,\omega)_*\mu}{\mathrm{d}\mu} \right\|_{L^2(\mu)} \le K \tag{46}$$

for all n and ω. Moreover, for every bounded measurable $g : X \to \mathbb{R}$ and $\varepsilon > 0$ there exists $N \ge 0$ such that the memory-loss property

$$\left| \int g(\varphi(n+m,\omega)x)\,\mathrm{d}\mu(x) - \int g(\varphi(n,\tau^m\omega)x)\,\mathrm{d}\mu(x) \right| < \varepsilon \tag{47}$$

hold for $n \ge N$, $m \ge 0$ and all ω. ■

The rest of the section is devoted to investigating some consequences of (SA5′).

Note that (47) asks that the integrals of $x \mapsto g((n,\tau^m\omega)x)$ with respect to the two measures $\varphi(m,\omega)_*\mu$ and μ are essentially the same for large n, uniformly in m and ω. The role of (46) is to allow for uniform approximations of the compositions $h \circ (\Phi^{(2)})^n$, $n \ge 0$, by compositions $\hat{h} \circ (\Phi^{(2)})^n$, where h is measurable and $\hat{h}$ is "simple": observe that $(h - \hat{h}) \circ (\Phi^{(2)})^n$ is not guaranteed to be uniformly (in n) small in $L^1(\bar{\mathbb{P}} \otimes \mu \otimes \mu)$, even if $h - \hat{h}$ is small, without some assumption. To that end, let us already prove a little lemma:

Lemma D.1 *Let $h : \Omega \times X \times X \to \mathbb{R}$ belong to $L^2(\bar{\mathbb{P}} \otimes \mu \otimes \mu)$. Then*

$$\|h \circ (\Phi^{(2)})^n\|_{L^1(\bar{\mathbb{P}}\otimes\mu\otimes\mu)} \le K^2 \|h\|_{L^2(\bar{\mathbb{P}}\otimes\mu\otimes\mu)}$$

holds for all $n \ge 0$ with K as in (46).

Proof Write $\lambda = \bar{\mathbb{P}} \otimes \mu \otimes \mu$ for brevity. Observe that

$$\begin{aligned}
&\left(\int |h| \circ (\Phi^{(2)})^n \, \mathrm{d}\lambda \right)^2 \\
&= \left(\int |h|(\tau^n \omega, x, y) \, \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(x) \, \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(y) \, \mathrm{d}\lambda(\omega, x, y) \right)^2 \\
&\le \int |h|^2(\tau^n \omega, x, y) \, \mathrm{d}\lambda(\omega, x, y) \int \left| \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(x) \, \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(y) \right|^2 \, \mathrm{d}\lambda(\omega, x, y)
\end{aligned}$$

by Hölder's inequality. Here

$$\int |h|^2(\tau^n \omega, x, y) \, \mathrm{d}\lambda(\omega, x, y) = \int |h|^2(\omega, x, y) \, \mathrm{d}\lambda(\omega, x, y)$$

since $\bar{\mathbb{P}}$ is stationary. On the other hand,

$$\begin{aligned}
&\int \left| \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(x) \, \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(y) \right|^2 \, \mathrm{d}\lambda(\omega, x, y) \\
&= \int \left[\int \left| \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(x) \right|^2 \mathrm{d}\mu(x) \int \left| \frac{\mathrm{d}\varphi(n,\omega)_* \mu}{\mathrm{d}\mu}(y) \right|^2 \mathrm{d}\mu(y) \right] \mathrm{d}\bar{\mathbb{P}}(\omega) \\
&\le K^4
\end{aligned}$$

by (46). Combining the estimates and taking square roots yields the result. □

D.1 Standing Assumption (SA5′) Implies (SA5)

Lemma D.2 *There exists an invariant measure* $\mathbf{P}^{(2)}$ *for the RDS* $\varphi^{(2)}$ *such that*

$$\lim_{n\to\infty} \int h \circ (\Phi^{(2)})^n \, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) = \int h \, \mathrm{d}\mathbf{P}^{(2)}$$

for all bounded measurable $h : \Omega \times X \times X \to \mathbb{R}$. *Moreover,* (30) *holds, and* $\mathbf{P}$ *in* (31) *is an invariant measure for the RDS* φ *such that*

$$\lim_{n\to\infty} \int \tilde{h} \circ \Phi^n \, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu) = \int \tilde{h} \, \mathrm{d}\mathbf{P}$$

for all bounded measurable $\tilde{h} : \Omega \times X \to \mathbb{R}$.

Proof Let $u : \Omega \to \mathbb{R}$ and $g^1, g^2 : X \to \mathbb{R}$ be bounded measurable. Let $\varepsilon > 0$. Then there exists $N \ge 0$ such that

$$\begin{aligned}
&\iiint (u \otimes g^1 \otimes g^2) \circ (\Phi^{(2)})^{n+m}(\omega, x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) \, \mathrm{d}\bar{\mathbb{P}}(\omega) \\
&= \int u(\tau^{n+m}\omega) \int g^1(\varphi(n+m,\omega)x) \, \mathrm{d}\mu(x) \int g^2(\varphi(n+m,\omega)y) \, \mathrm{d}\mu(y) \, \mathrm{d}\bar{\mathbb{P}}(\omega) \\
&= \int u(\tau^{n+m}\omega) \int g^1(\varphi(n,\tau^m\omega)x) \, \mathrm{d}\mu(x) \int g^2(\varphi(n,\tau^m\omega)y) \, \mathrm{d}\mu(y) \, \mathrm{d}\bar{\mathbb{P}}(\omega) + O(\varepsilon) \\
&= \int u(\tau^{n}\omega) \int g^1(\varphi(n,\omega)x) \, \mathrm{d}\mu(x) \int g^2(\varphi(n,\omega)y) \, \mathrm{d}\mu(y) \, \mathrm{d}\bar{\mathbb{P}}(\omega) + O(\varepsilon) \\
&= \iiint (u \otimes g^1 \otimes g^2) \circ (\Phi^{(2)})^{n}(\omega, x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) \, \mathrm{d}\bar{\mathbb{P}}(\omega) + O(\varepsilon)
\end{aligned}$$

for all $n \geq N$ and $m \geq 0$. Here the third line uses (47) and the fourth line uses stationarity. Thus, we see that the sequence $\left(\iiint (u \otimes g^1 \otimes g^2) \circ (\Phi^{(2)})^n(\omega, x, y)\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)\, \mathrm{d}\bar{\mathbb{P}}(\omega)\right)_n$ is Cauchy and therefore convergent. We will show using the monotone class theorem that the convergence property extends to an arbitrary bounded measurable function in place of $u \otimes g^1 \otimes g^2$.

Let $\mathcal{H}$ denote the set of all measurable functions $h : \Omega \times X \times X \to \mathbb{R}$ such that $\lim_{n\to\infty} \int h \circ (\Phi^{(2)})^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu)$ exists. Let $\mathcal{A}$ denote the set of all measurable cubes in $\Omega \times X \times X$. Clearly $\mathcal{A}$ is nonempty and closed under finite intersections, and it contains the product space $\Omega \times X \times X$. Clearly $\mathcal{H}$ is closed under linear combinations. Furthermore, the argument above shows $1_A \in \mathcal{H}$ for all $A \in \mathcal{A}$. Suppose now that $h_k \in \mathcal{H}$ are nonnegative functions increasing to a bounded function h. Showing $h \in \mathcal{H}$ proves that $\mathcal{H}$ contains all bounded functions that are measurable with respect to the sigma-algebra $\sigma(\mathcal{A}) = \mathcal{F} \otimes \mathcal{B} \otimes \mathcal{B}$. We will show $h \in \mathcal{H}$ next.

Let $\varepsilon > 0$ be fixed. Since $0 \leq h_k \uparrow h$ where h is bounded, by the bounded convergence theorem there exists $k_0 = k_0(\varepsilon)$ such that $\|h - h_{k_0}\|_{L^2(\bar{\mathbb{P}}\otimes\mu\otimes\mu)} < \varepsilon$. Thus, by Lemma D.1,

$$\|(h - h_{k_0}) \circ (\Phi^{(2)})^n\|_{L^1(\bar{\mathbb{P}}\otimes\mu\otimes\mu)} < K^2\varepsilon$$

for all $n \geq 1$. Since $h_{k_0} \in \mathcal{H}$, there exists $n_0 = n_0(\varepsilon)$ such that

$$\left| \int h_{k_0} \circ (\Phi^{(2)})^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) - \lim_{m\to\infty} \int h_{k_0} \circ (\Phi^{(2)})^m\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) \right| < \varepsilon$$

for all $n \geq n_0$. A combination of the estimates yields

$$\left| \int h \circ (\Phi^{(2)})^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) - \lim_{m\to\infty} \int h_{k_0} \circ (\Phi^{(2)})^m\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) \right| < K^2\varepsilon + \varepsilon$$

for all $n \geq n_0$. Hence $h \in \mathcal{H}$. Therefore, by the monotone class theorem $\mathcal{H}$ contains all bounded measurable functions.

By the Vitali–Hahn–Saks theorem there exists a probability measure $\mathbf{P}^{(2)}$ satisfying

$$\lim_{n\to\infty} \int h \circ (\Phi^{(2)})^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) = \int h\, \mathrm{d}\mathbf{P}^{(2)}$$

for all bounded measurable $h : \Omega \times X \times X \to \mathbb{R}$. The symmetry property (30) of $\mathbf{P}^{(2)}$ is an immediate consequence. By the same token $\mathbf{P}^{(2)}$ is invariant for $\Phi^{(2)}$:

$$\int h \circ \Phi^{(2)}\, \mathrm{d}\mathbf{P}^{(2)} = \lim_{n\to\infty} \int h \circ \Phi^{(2)} \circ (\Phi^{(2)})^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) = \int h\, \mathrm{d}\mathbf{P}^{(2)}.$$

Furthermore, taking h of the form $h(\omega, x, y) = u(\omega)$,

$$\int h\, \mathrm{d}\mathbf{P}^{(2)} = \lim_{n\to\infty} \int u(\tau^n\omega)\, \mathrm{d}\bar{\mathbb{P}}(\omega) = \int u\, \mathrm{d}\bar{\mathbb{P}}$$

shows $(\Pi_1)_*\mathbf{P}^{(2)} = \bar{\mathbb{P}}$. Thus, $\mathbf{P}^{(2)}$ is an invariant measure for the RDS $\varphi^{(2)}$.

Suppose that either $h(\omega, x, y) = \tilde{h}(\omega, x)$ or $h(\omega, x, y) = \tilde{h}(\omega, y)$ holds identically. Then

$$\int \tilde{h}\, \mathrm{d}\mathbf{P} = \int h\, \mathrm{d}\mathbf{P}^{(2)} = \lim_{n\to\infty} \int h \circ (\Phi^{(2)})^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) = \lim_{n\to\infty} \int \tilde{h} \circ \Phi^n\, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu).$$

This yields the claims concerning $\mathbf{P}$. □

We are in position to prove the promised fact:

Lemma D.3 *Standing Assumption* (SA5′) *implies* (SA5).

Proof By Lemma D.2 it remains to verify (32)–(34). Using Lemma D.2,

$$\lim_{i\to\infty} \bar{\mathbb{E}}[\mu(f_i)] = \lim_{i\to\infty} \int f \circ \Pi_2 \circ \Phi^i \, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu) = \int f \circ \Pi_2 \, \mathrm{d}\mathbf{P}$$

and

$$\lim_{i\to\infty} \bar{\mathbb{E}}[\mu(f_i f_{i+k})] = \lim_{i\to\infty} \int (f \circ \Pi_2 \; f \circ \Pi_2 \circ \Phi^k) \circ \Phi^i \, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu) = \int f \circ \Pi_2 \; f \circ \Pi_2 \circ \Phi^k \, \mathrm{d}\mathbf{P}.$$

Likewise

$$\begin{aligned} \lim_{i\to\infty} \bar{\mathbb{E}}[\mu(f_i)\mu(f_{i+k})] &= \lim_{i\to\infty} \int (f \circ \Pi_2 \; f \circ \Pi_3 \circ (\Phi^{(2)})^k) \circ (\Phi^{(2)})^i \, \mathrm{d}(\bar{\mathbb{P}} \otimes \mu \otimes \mu) \\ &= \int f \circ \Pi_2 \; f \circ \Pi_3 \circ (\Phi^{(2)})^k \, \mathrm{d}\mathbf{P}^{(2)}. \end{aligned}$$

The proof is complete. □

D.2 Disintegration of the Invariant Measure $\mathbf{P}^{(2)}$

In this subsection we shed some light on the invariant measure $\mathbf{P}^{(2)}$ of the RDS $\varphi^{(2)}$ with the aid of disintegrations. The mathematical constructions here are well known, and we include this part for completeness. The results call for nice structure of the measurable spaces: we assume that both $(X, \mathcal{B})$ and $(\Omega_0, \mathcal{E})$ are **standard measurable spaces**.

We begin by stating a basic fact:

Lemma D.4 *There exists a family of set functions* $\nu_\omega^{(2)} : \mathcal{B} \to [0, 1]$, $\omega \in \Omega$, *such that*

(i) *the map* $\omega \mapsto \nu_\omega^{(2)}(B)$ *is measurable for all* $B \in \mathcal{B} \otimes \mathcal{B}$;
(ii) $\nu_\omega^{(2)}$ *is a probability measure for* $\bar{\mathbb{P}}$*-a.e.* $\omega \in \Omega$;
(iii) *for all* $h \in L^1(\mathbf{P}^{(2)})$,

$$\int h \, \mathrm{d}\mathbf{P}^{(2)} = \int_\Omega \int_{X\times X} h(\omega, x, y) \, \mathrm{d}\nu_\omega^{(2)}(x, y) \, \mathrm{d}\bar{\mathbb{P}}(\omega).$$

The disintegration is essentially unique: if $\tilde{\nu}_\omega^{(2)}$, $\omega \in \Omega$, *is another family of such set functions, then* $\nu_\omega^{(2)} = \tilde{\nu}_\omega^{(2)}$ *for* $\bar{\mathbb{P}}$*-a.e.* $\omega \in \Omega$.

Proof Since the product space $(X \times X, \mathcal{B} \otimes \mathcal{B})$ is also a standard measurable space and $(\Pi_1)_*\mathbf{P} = \bar{\mathbb{P}}$, classical results yield the lemma; see, e.g., Arnold [Proposition 1.4.3][3]. □

It is helpful to think of $\nu_\omega^{(2)}$ as the conditional measure $\mathbf{P}^{(2)}(\,\cdot \mid \omega)$. In the following we will characterize the conditional measures $\nu_\omega^{(2)}$.

Next, we extend $\bar{\mathbb{P}}$ to a stationary measure on the space of two-sided sequences. To that end define $\Omega^- = \Omega_0^{\{\dots,-2,-1,0\}}$ and $\Omega^+ = \Omega_0^{\{1,2,3,\dots\}} = \Omega$. The sigma-algebras $\mathcal{F}^-$ and $\mathcal{F}^+ = \mathcal{F}$ denote the corresponding products of $\mathcal{E}$. Write also

$$\bar{\Omega} = \Omega^- \times \Omega^+ = \Omega_0^{\mathbb{Z}} \quad \text{and} \quad \bar{\mathcal{F}} = \mathcal{F}^- \otimes \mathcal{F}^+ = \mathcal{E}^{\mathbb{Z}}.$$

Let $\bar{\tau} : \bar{\Omega} \to \bar{\Omega}$ denote the two-sided shift: $(\bar{\tau}^k \bar{\omega})_i = \bar{\omega}_{i+k}$ for all $i, k \in \mathbb{Z}$. Finally, let $\Pi^\pm : \bar{\Omega} \to \Omega^\pm$ denote the canonical projections: $\Pi^-(\bar{\omega}) = \omega^-$ and $\Pi^+(\bar{\omega}) = \omega^+$ for all $\bar{\omega} = (\omega^-, \omega^+) \in \Omega^- \times \Omega^+$.

We are ready to state another basic fact:

Lemma D.5 (1) *There exists a unique probability measure $\bar{\mathbb{Q}}$ on $(\bar{\Omega}, \bar{\mathcal{F}})$ which is invariant for $\bar{\tau}$ and satisfies $(\Pi^+)_*\bar{\mathbb{Q}} = \bar{\mathbb{P}}$.*
(2) *There exists an essentially unique family of set functions $q_\omega : \mathcal{F}^- \to [0, 1]$, $\omega \in \Omega$, such that*

(i) *the map $\omega \mapsto q_\omega(E)$ is measurable for all $E \in \mathcal{F}^-$;*
(ii) *q_ω is a probability measure for $\bar{\mathbb{P}}$-a.e. $\omega \in \Omega$;*
(iii) *for all $h \in L^1(\bar{\mathbb{Q}})$,*

$$\int_{\bar{\Omega}} h(\bar{\omega})\, \mathrm{d}\bar{\mathbb{Q}}(\bar{\omega}) = \int_{\Omega}\int_{\Omega^-} h(\omega^-, \omega)\, \mathrm{d}q_\omega(\omega^-)\, \mathrm{d}\bar{\mathbb{P}}(\omega).$$

Proof (1) Since $(\Omega_0, \mathcal{E})$ is a standard measurable space, the shift-invariant measure $\bar{\mathbb{Q}}$ having $\bar{\mathbb{P}}$ as its marginal is uniquely constructed with the aid of Kolmogorov's extension theorem by requiring that the finite dimensional distributions are translation invariant and coincide with those of $\bar{\mathbb{P}}$. See, e.g., Arnold [3, Appendix A.3] for details.

(2) Since $(\Omega^-, \mathcal{F}^-)$ is a standard probability space and $(\Pi^+)_*\bar{\mathbb{Q}} = \bar{\mathbb{P}}$, the result is classical as in Lemma D.4. □

The resulting dynamical system $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{Q}}, \bar{\tau})$ is the natural extension of $(\Omega, \mathcal{F}, \bar{\mathbb{P}}, \tau)$ with homomorphism Π^+. The intuition behind the measures in Lemma D.5 is the following: Think of $\omega = (\omega_1, \omega_2, \ldots)$ as a stochastic process with law $\bar{\mathbb{P}}$. Due to stationarity, it is possible to glue a history $\omega^- = (\ldots, \omega_{-1}, \omega_0)$ to ω in a consistent and unique way such that the law $\bar{\mathbb{Q}}$ of $\bar{\omega} = (\omega^-, \omega) = (\ldots, \omega_{-1}, \omega_0, \omega_1, \omega_2, \ldots)$ is stationary and the marginal law corresponding to the future part ω is $\bar{\mathbb{P}}$. The measure q_ω can be thought of as the conditional law $\bar{\mathbb{Q}}(\,\cdot\,|\,\omega)$, the distribution of the past ω^- given the future ω.

For the following it will be convenient to introduce the notations

$$\varphi(\omega^-_{-n+1}, \ldots, \omega^-_0) = T_{\omega^-_0} \circ \cdots \circ T_{\omega^-_{-n+1}}$$

and

$$\varphi^{(2)}(\omega^-_{-n+1}, \ldots, \omega^-_0)(x, y) = (\varphi(\omega^-_{-n+1}, \ldots, \omega^-_0)x,\ \varphi(\omega^-_{-n+1}, \ldots, \omega^-_0)y)$$

for any finite sequence $(\omega^-_{-n+1}, \ldots, \omega^-_0) \subset \Omega_0$.

Now, for all bounded measurable functions $h(\omega, x, y) = u(\omega)g(x, y)$ we have

$$\int h\, \mathrm{d}\mathbf{P}^{(2)} = \int_{\Omega} u(\omega) \int_{X\times X} g\, \mathrm{d}\nu^{(2)}_\omega\, \mathrm{d}\bar{\mathbb{P}}(\omega). \tag{48}$$

On the other hand, Lemma D.2 yields

$$\begin{aligned}
\int h\, \mathrm{d}\mathbf{P}^{(2)} &= \lim_{n\to\infty} \int_{\Omega} u(\tau^n\omega) \int_{X\times X} g \circ \varphi^{(2)}(n, \omega)\, \mathrm{d}(\mu\otimes\mu)\, \mathrm{d}\bar{\mathbb{P}}(\omega) \\
&= \lim_{n\to\infty} \int_{\bar{\Omega}} u(\Pi^+(\bar{\tau}^n\bar{\omega})) \int_{X\times X} g \circ \varphi^{(2)}(n, \Pi^+(\bar{\omega}))\, \mathrm{d}(\mu\otimes\mu)\, \mathrm{d}\bar{\mathbb{Q}}(\bar{\omega}) \\
&= \lim_{n\to\infty} \int_{\bar{\Omega}} u(\Pi^+(\bar{\omega})) \int_{X\times X} g \circ \varphi^{(2)}(n, \Pi^+(\bar{\tau}^{-n}\bar{\omega}))\, \mathrm{d}(\mu\otimes\mu)\, \mathrm{d}\bar{\mathbb{Q}}(\bar{\omega}).
\end{aligned}$$

In order to disintegrate $\bar{\mathbb{Q}}$, let us write $\bar{\omega} = (\omega^-, \omega)$ in the obvious manner, noting that $u(\Pi^+(\bar{\omega})) = u(\omega)$ and $\varphi^{(2)}(n, \Pi^+(\bar{\tau}^{-n}\bar{\omega})) = \varphi^{(2)}(\omega^-_{-n+1}, \ldots, \omega^-_0)$. Thus, Lemma D.5 yields

$$\int h\, \mathrm{d}\mathbf{P}^{(2)} = \lim_{n\to\infty} \int_{\Omega} u(\omega) \int_{\Omega^-}\int_{X\times X} g \circ \varphi^{(2)}(\omega^-_{-n+1}, \ldots, \omega^-_0)\, \mathrm{d}(\mu\otimes\mu)\, \mathrm{d}q_\omega(\omega^-)\, \mathrm{d}\bar{\mathbb{P}}(\omega). \tag{49}$$

The following observation is now key:

Lemma D.6 *Given $\omega^- \in \Omega^-$, there exists a probability measure μ_{ω^-} on $(X, \mathcal{B})$ such that*

$$\lim_{n\to\infty} \int_{X\times X} g \circ \varphi^{(2)}(\omega^-_{-n+1}, \dots, \omega^-_0)\, \mathrm{d}(\mu \otimes \mu) = \int_{X\times X} g\, \mathrm{d}(\mu_{\omega^-} \otimes \mu_{\omega^-}) \tag{50}$$

for all bounded measurable $g : X \times X \to \mathbb{R}$.

Note that μ_{ω^-} has the interpretation of being the pushforward of μ from the infinitely distant past along the history $\omega^- = (\dots, \omega^-_1, \omega^-_0)$.

Proof of Lemma D.6 Consider first a bounded measurable $g^1 : X \to \mathbb{R}$. Let $\varepsilon > 0$. By (47) of (SA5′) there exists $N \geq 0$ such that

$$\left| \int g^1 \circ \varphi(\omega^-_{-n-m+1}, \dots, \omega^-_0)\, \mathrm{d}\mu - \int g^1 \circ \varphi(\omega^-_{-n+1}, \dots, \omega^-_0)\, \mathrm{d}\mu \right| < \varepsilon \tag{51}$$

for all $n \geq N$ and $m \geq 0$. We see that $(\int g^1 \circ \varphi(\omega^-_{-n+1}, \dots, \omega^-_0)\, \mathrm{d}\mu)_{n=1}^{\infty}$ is a Cauchy sequence, and thus convergences. Since g^1 was arbitrary, the Vitali–Hahn–Saks theorem yields the existence of a measure μ_{ω^-} such that

$$\lim_{n\to\infty} \int g^1 \circ \varphi(\omega^-_{-n+1}, \dots, \omega^-_0)\, \mathrm{d}\mu = \int g^1\, \mathrm{d}\mu_{\omega^-}.$$

This yields (50) for all $g(x, y) = g^1(x)g^2(y)$ with both $g^1, g^2 : X \to \mathbb{R}$ bounded and measurable. Similarly to the proof of Lemma D.2, a straightforward application of the monotone class theorem extends (50) to all bounded measurable $g : X \times X \to \mathbb{R}$. □

We finally arrive at the characterization of the conditional measure $\nu^{(2)}_\omega$ as the expected pushforward of $\mu \otimes \mu$ from the infinitely distant past along all histories consistent with ω:

Corollary D.7 *For $\bar{\mathbb{P}}$-a.e. $\omega \in \Omega$,*

$$\nu^{(2)}_\omega(\,\cdot\,) = \int_{\Omega^-} (\mu_{\omega^-} \otimes \mu_{\omega^-})(\,\cdot\,)\, \mathrm{d}q_\omega(\omega^-).$$

Proof Equating first the expressions of $\int h\, \mathrm{d}\mathbf{P}^{(2)}$ in (48) and (49), and then applying (50) to the latter, we obtain

$$\int_\Omega u(\omega) \int_{X\times X} g\, \mathrm{d}\nu^{(2)}_\omega\, \mathrm{d}\bar{\mathbb{P}}(\omega) = \int_\Omega u(\omega) \int_{\Omega^-} \int_{X\times X} g\, \mathrm{d}(\mu_{\omega^-} \otimes \mu_{\omega^-})\, \mathrm{d}q_\omega(\omega^-)\, \mathrm{d}\bar{\mathbb{P}}(\omega).$$

Since the conditional measures $\nu^{(2)}_\omega$ are unique, the claim follows. □

Let us lastly point out that the invariance of $\mathbf{P}^{(2)}$ is equivalent to

$$\bar{\mathbb{E}}[\varphi^{(2)}(m, \,\cdot\,)_* \nu^{(2)}_{(\cdot)} \,|\, \tau^{-m}\mathcal{F}](\omega) = \nu^{(2)}_{\tau^m\omega}$$

holding for almost all ω with respect to $\bar{\mathbb{P}}$, for all $m \geq 1$. The equation means that

$$\int_\Omega \varphi^{(2)}(m, \omega)_* \nu^{(2)}_\omega(g)\, u(\tau^m\omega)\, \mathrm{d}\bar{\mathbb{P}}(\omega) = \int_\Omega \nu^{(2)}_{\tau^m\omega}(g)\, u(\tau^m\omega)\, \mathrm{d}\bar{\mathbb{P}}(\omega)$$

holds for all bounded measurable functions $g : X \times X \to \mathbb{R}$ and $u : \Omega \to \mathbb{R}$. It is a good exercise for the interested reader to reprove the invariance of $\mathbf{P}^{(2)}$ by verifying the equation above directly, using Corollary D.7, Lemma D.6 and Lemma D.5.

References

1. Abdelkader, M., Aimino, R.: On the quenched central limit theorem for random dynamical systems. J. Phys. A. https://doi.org/10.1088/1751-8113/49/24/244002 (2016)
2. Aimino, R., Nicol, M., Vaienti, S.: Annealed and quenched limit theorems for random expanding dynamical systems. Probab. Theory Relat. Fields **162**(1–2), 233–274 (2015). https://doi.org/10.1007/s00440-014-0571-y
3. Arnold, L.: Random Dynamical Systems. Springer Monographs in Mathematics. Springer, Berlin (1998). https://doi.org/10.1007/978-3-662-12878-7
4. Ayyer, A., Liverani, C., Stenlund, M.: Quenched CLT for random toral automorphism. Discrete Contin. Dyn. Syst. **24**(2), 331–348 (2009). https://doi.org/10.3934/dcds.2009.24.331
5. Bahsoun, W., Bose, C.: Corrigendum: mixing rates and limit theorems for random intermittent maps (2016 Nonlinearity 29 1417). Nonlinearity **29**(12), C4 (2016). https://doi.org/10.1088/0951-7715/29/12/C4
6. Bahsoun, W., Bose, C., Duan, Y.: Decay of correlation for random intermittent maps. Nonlinearity **27**(7), 1543–1554 (2014). https://doi.org/10.1088/0951-7715/27/7/1543
7. Bahsoun, W., Bose, C., Ruziboev, M.: Quenched decay of correlations for slowly mixing systems. arXiv:1706.04158 (2017)
8. Chen, J., Yang, Y., Zhang, H.-K.: Non-stationary almost sure invariance principle for hyperbolic systems with singularities. J. Stat. Phys. (2018). https://doi.org/10.1007/s10955-018-2107-9
9. Dragičević, D., Froyland, G., Gonzalez-Tokman, C., Vaienti, S.: A spectral approach for quenched limit theorems for random expanding dynamical systems. Commun. Math. Phys. **360**(3), 1121–1187 (2018). https://doi.org/10.1007/s00220-017-3083-7
10. Dragičević, D., Froyland, G., González-Tokman, C., Vaienti, S.: Almost sure invariance principle for random piecewise expanding maps. Nonlinearity **31**(5), 2252–2280 (2018). https://doi.org/10.1088/1361-6544/aaaf4b
11. Freitas, A.C.M., Freitas, J.M., Vaienti, S.: Extreme value laws for non stationary processes generated by sequential and random dynamical systems. Ann. Inst. Henri Poincaré Probab. Stat. **53**(3), 1341–1370 (2017). https://doi.org/10.1214/16-AIHP757
12. Gál, I.S., Koksma, J.F.: Sur l'ordre de grandeur des fonctions sommables. Nederl. Akad. Wetensch., Proc. **53**, 638–653 = Indagationes Math. **12**, 192–207 (1950)
13. Hella, O.: Central limit theorems with a rate of convergence for sequences of transformations. Preprint. arXiv:1811.06062 (2018)
14. Kifer, Y.: Limit theorems for random transformations and processes in random environments. Trans. Am. Math. Soc. **350**(4), 1481–1518 (1998). https://doi.org/10.1090/S0002-9947-98-02068-6
15. Leonov, V.P.: On the dispersion of time means of a stationary stochastic process. Teor. Verojatnost. i Primen. **6**, 93–101 (1961)
16. Leppänen, J., Stenlund, M.: Sunklodas' approach to normal approximation for time-dependent dynamical systems. Preprint. arXiv:1906.03217 (2019)
17. Melbourne, I., Nicol, M.: A vector-valued almost sure invariance principle for hyperbolic dynamical systems. Ann. Probab. **37**(2), 478–505 (2009). https://doi.org/10.1214/08-AOP410
18. Nándori, P., Szász, D., Varjú, T.: A central limit theorem for time-dependent dynamical systems. J. Stat. Phys. **146**(6), 1213–1220 (2012). https://doi.org/10.1007/s10955-012-0451-8
19. Nicol, M., Török, A., Vaienti, S.: Central limit theorems for sequential and random intermittent dynamical systems. Ergod. Theory Dyn. Syst. **38**(3), 1127–1153 (2018). https://doi.org/10.1017/etds.2016.69
20. Philipp, W., Stout, W.: Almost sure invariance principles for partial sums of weakly dependent random variables. Mem. Am. Math. Soc. 2 (issue 2, 161), iv+140 (1975)